Contacts by distressed individuals to UK parapsychology and anomalous experience academic research units---A retrospective survey looking to the future

European Journal of Parapsychology

Volume 23.2, 2008

European Journal of Parapsychology

Volume 23.2 2008

European Journal of Parapsychology

2008
Volume 23.2

European Journal of Parapsychology

Psychology (DEC), Poole House, Bournemouth University, Poole BH12 5BB, UK
Indexed by PsycINFO.
Printed by JEM Digital Print Services Ltd. on recycled paper.

European Journal of Parapsychology (ISSN: 0168-7263) is published twice a year in Spring and Autumn. Subscription enquiries to subscriptions@ejp.org.uk or postal mail to the Editor at the above address.

INSTRUCTIONS TO AUTHORS

The European Journal of Parapsychology (EJP) is a peer-reviewed scientific journal for research - particularly theoretical and theory-driven empirical work - relating to the field of parapsychology (defined as the study of communication or interaction between organisms and their environment that do not appear to rely on the established sensorimotor channels). The Journal's aim is to stimulate and enhance activity in parapsychology, particularly in Europe, by publishing Articles, Research Notes, Reviews and Comments that offer insight into or criticisms of parapsychological research. The Journal also publishes peer-reviewed Student Research Briefs. Submissions are welcomed on any topic that falls under the heading of parapsychology as defined above. Authors are requested to briefly define any specific terms used (including 'standard' terms such as ESP, PK) in the text of the paper. All manuscripts submitted to the EJP must be under consideration solely by this journal and may not have been previously published in any form (with the exception of translations of non-English articles). Initial submissions should be sent via email to submissions@ejp.org.uk as an attached file (acceptable formats are ASCII, PDF, RTF, OpenOffice or MS Word).

ARTICLES are reports of experimental, clinical or theoretical research, with introductory comments and discussion of results generally being of direct relevance to the data or ideas being presented. Generally articles should be less than 10,000 words in length, though longer papers may be accepted after prior negotiation with the Editors. After evaluation by the editors, the initial article will be sent to two referees for peer-review.

RESEARCH NOTES are manuscripts of up to 3000 words in the following categories: conceptual replications; attempted replications which contradict the original study; reporting on new methodology or apparatus; new applications of existing methodology; challenges to key elements of an existing study, theory or methodology. Research Notes are not intended as a way of publishing inadequate or incomplete work. The Editors or Referees may also decide that a manuscript submitted as an article would be better published as a Research Note.

STUDENT RESEARCH BRIEFS are summaries of research (maximum 2000 words)conducted for assessment as part of a Bachelors or Masters degree (or equivalent) which would not otherwise be submitted as a full article but where the research is of sufficient interest to be disseminated to the wider parapsychological community. Student Research Briefs should provide a summary of methodology and results of the conducted research. They will be subject to peer-review with the understanding that the scope of the research will be less extensive than that of a full article.

REVIEWS provide critical and analytical review or assessment of previously published work. The EJP welcomes novel and cogent reviews (1500 words maximum), subject to Editorial approval.

COMMENTS are extended letters (1000 words maximum) that address specific issues raised in material published in the journal. All comments will be sent for peer-review, the process being similar to article review. Comments will be published in as timely a manner as possible, accompanied where appropriate with a reply from the original author(s).

Full formatting details are provided on the EJP website http://ejp.org.uk/

European Journal of Parapsychology

Volume 23.2

2008

European Journal of Parapsychology
Volume 23.2, pages 131-146

ISSN: 0168-7263

Correlations between the EEGs of two spatially separated subjects – a replication study

Wolfgang Ambach
Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany

Abstract

Previous studies have reported correlations between the EEGs of two spatially separated subjects. These correlations could not be explained by means of conventional physiological or physical mechanisms. Results by Wackermann (2003, 2004; Wackermann et al, 2004) suggest such unexplained correlations, with the direction (increase or decrease of EEG power) as well as the localization (electrode position) of the supposed effects varying between studies. The present replication study, performed in another laboratory, yet other experimental changes kept at a minimum, aimed at clarifying this effect. Seventeen pairs of subjects were examined. One subject of each pair was visually stimulated by one-second periods of checkerboard reversals while the other sat in a dimly lit separate room without stimulation. The latter subject's EEG power in the paired subject's stimulation and interstimulus periods was compared; separate analyses were performed for two different time windows. In this study, the previously used mode of EEG segment sampling turned out to systematically overestimate effects. Therefore, the non-parametric statistical method was modified. Also, the results for the nineteen EEG channels were corrected for their degrees of freedom using a Monte-Carlo simulation. Results did not show a significant deviation from random expectation; the EEG power of the non-stimulated subject was not linked to the stimulation of the other subject.

Correspondence details: Wolfgang Ambach, Institut für Grenzgebiete der Psychologie und Psychohygiene, Wilhelmstr. 3A, D-79098 Freiburg, Germany. Email: ambach@igpp.de

Introduction

Since 1965, a series of studies have reported correlations between the EEGs of spatially separated subjects. In these experiments, pairs of subjects were prevented from interacting via known physiological or physical mechanisms; consequently, the findings have often been discussed in terms of direct brain-to-brain communication (Charman, 2006).

Grinberg-Zylberbaum et al. (1994) coined the term 'transferred potentials' for the observation of correlations between the EEGs of two spatially separated subjects of whom only one was visually stimulated. They found statistical correlations between certain parts of the two time series of the averaged event-related EEG segments obtained from both subjects. From there on, event-related EEG correlations between distant human subjects were mostly investigated on the basis of averaged EEG segments. A critical analysis of the Grinberg-Zylberbaum et al. (1994) study has been provided by May et al. (2001). They provided evidence for the claim that violations of the underlying assumptions concerning hypothesis testing had led to an overestimation of the effects in this study. Several subsequent experiments searching for EEG correlations which could not be explained by a known mechanism, yielded positive results (e.g. Fenwick et al., 1998), while other replication studies (e.g. Sabell et al., 2001) failed to confirm the effect. Recently, correlations between brain electrical activities of distant human brains were found by Standish (2004) and Radin (2003, 2004).

Two studies by Wackermann et al. (2003, 2004) aimed at examining if correlations in the EEG between a visually stimulated subject and a second, non-stimulated subject were replicable when communication between both participants was prevented by spatial separation of subjects in shielded rooms. One subject ('subject A') was visually stimulated with repeated series of checkerboard reversals while the other subject ('subject B'), sitting in a dimly lit room, was not stimulated. In the control condition, stimulation of the first subject was prevented by covering the monitor with an opaque shield. From the first to the second study, several methodological refinements had been introduced. In both studies, the authors found significant fluctuations of the EEG power of subject B according to the stimulation or interstimulus periods of subject A. Remarkably, the effect of subject A's stimulation on the EEG power of the non-stimulated subject B varied between studies in direction (increase, decrease) as well as location (electrode position). Furthermore, deviations from random expectation were not only observed in the experimental condition (stimulation), but partially also in the control condition (no stimulation). Thus, significant results were generally replicated, but not the specific finding of the direction of the effect or the location of its maximum; results of the repeated studies were inconsistent but still striking.

Our replication study aimed at validating the existence of the effect and, if possible, shedding light upon its specifications. In relation to the preceding studies, it was conducted in a different laboratory and guided by a different experimenter. All other experimental conditions were kept unchanged as far as possible. The statistical methods applied in the precedent studies were reconsidered in detail.

Method

Subjects

Seventeen pairs of related subjects (friends, relatives, or partners; 12 male, 22 female; age 23.2 ± 1.9 years; mean duration of relationship 4.8 ± 5.3 years) were recruited via announcement in the local students' job agency. They were of reportedly good health, unmedicated, and voluntarily participated in the study for the payment of 17 Euros. Informed consent was obtained from all subjects.

Stimuli

As visual stimulus, a checkerboard pattern with 64 x 48 black and white squares (angular size ≈ 0.25° x 0.25°) was presented on a 19'' LCD monitor. The stimulated subject observed the monitor from ≈ 125 cm distance (visual angle 16° x 12°). Duration of each stimulation period was one second, consisting of four 250 ms periods with reversed pattern presentations (see Figure 1).

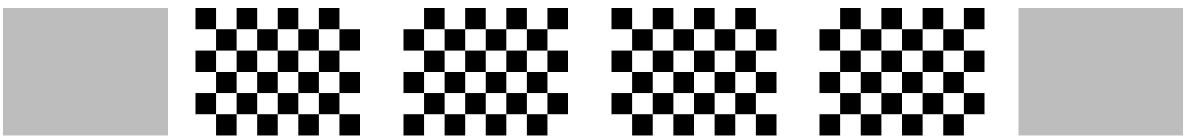


Figure 1. Visual stimulation using checkerboard reversals

168 stimuli were presented within each of the two runs of the experiment according to Wackermann et al. (2004). Interstimulus intervals were 1.5–7.5 seconds; the sequence of intervals was pseudo-randomized. During interstimulus intervals, the screen was uniformly grey with the brightness adjusted to the overall brightness of the checkerboard stimuli.

Physiological Recording

EEG data were recorded from nineteen electrodes according to the standard 10-20 positions (Jasper, 1958), against linked mastoids as reference. Impedances were kept below 5 kΩ; Ag-AgCl electrodes were attached to the scalp with use of Easy Caps and abrasive gel (both by Falk Minow, München, Germany). Vertical and horizontal oculogram was recorded for artifact control. Two Synamps amplifiers (NeuroScan, Inc.) of the same type were used. EEG was band-pass filtered at 0.15 - 70 Hz, A/D-converted, sampled at 1000 Hz and saved to disc using the software *Acquire*, Version 4.3.1 (Neurosoft, Inc.).

Procedure

During the experiment, the subjects were seated in two spatially separate chambers (distance: 8 meters). The experimental rooms were acoustically and electromagnetically shielded and located on different sides of the laboratory main room (Figure 2).

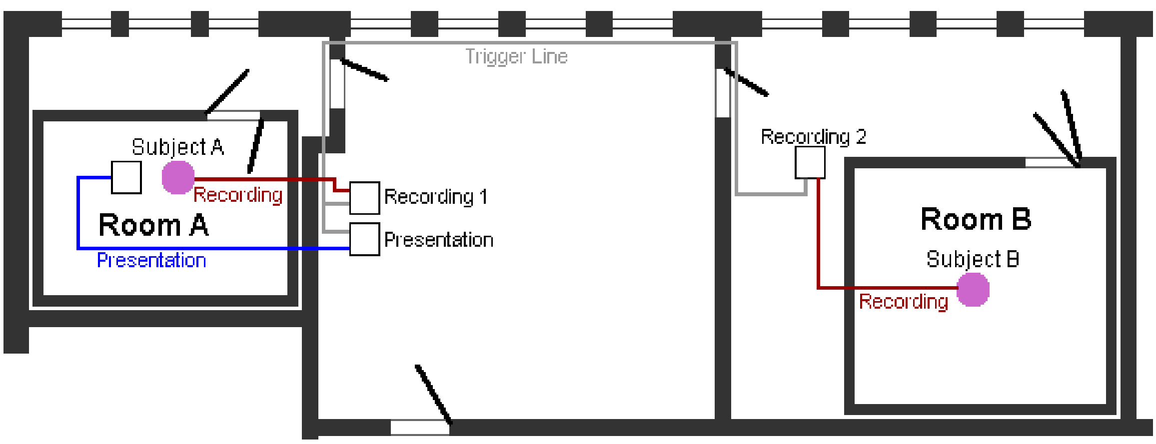


Figure 2. Spatial arrangement of the experimental setup

Each subject was observed via a video-camera by one of the two experimenters. While one subject of each pair, randomly assigned, was stimulated visually in room A, the second was instructed to relax with eyes open in the dimly lit room B. The whole experiment consisted of two runs, the 'uncovered' (experimental) and the 'covered' (control) condition; the sequence of both was balanced over the study. Between both runs, the experimenter responsible for the subject in room A entered the chamber and either covered or uncovered the monitor. The break between both runs was kept between three and five minutes. The non-stimulated subject was not informed about the change of conditions.

Data analysis

For the analysis, offline EEG data were filtered to 0.15 - 30 Hz and visually inspected for artifacts. Data from both stimulated (A) and non-stimulated (B) subjects were segmented into stimulation and interstimulus epochs of one second duration, according to subject A's stimulation procedure. Segments contaminated with artifacts were discarded from further evaluation; the average loss rate was 34%. For subsequent analyses, the valid segments were detrended by subtracting each segment's third-order polynomial approximation from the data.

In order to compare subject B's average EEG power between subject A's stimulation and interstimulus epochs, a distribution-free boot-strapping procedure was developed and applied

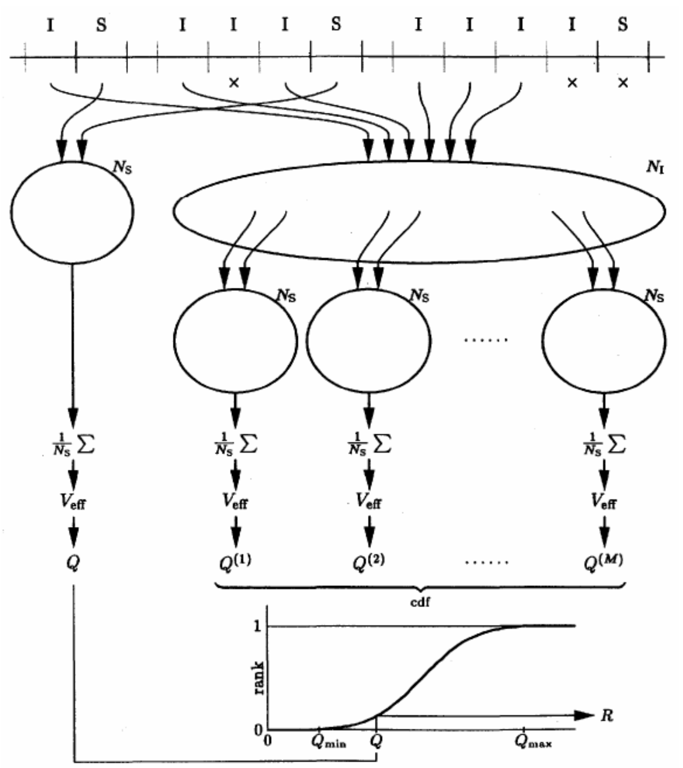


Figure 3. Flow chart of EEG data processing (from Wackermann et al., 2004)

by Wackermann et al (2004). Figure 3 illustrates the procedure as a flow chart (from Wackermann, 2004).

For each non-stimulated subject, valid stimulation (S) and interstimulus (I) EEG epochs were collected in two pools. All N_S stimulation epochs were averaged; the EEG power V_{eff} was calculated for the averaged signal as theroot mean square of voltage u_i of the n single data points: $V_{eff} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} u_i^{\,2}}$. From the N_I interstimulus epochs, 10000 subsets were drawn randomly, each comprising N_S segments. Within each subset, the EEG power was computed for the averaged signal. The rank R of the stimulation epochs EEG power was determined within the cumulative distribution function of the interstimulus EEG power values for the 10000 subsets. The procedure was performed separately for each non-stimulated subject and for each EEG channel. Ranks R, after division by the number of subsets, were within the interval [0,1] and reflected the probability of receiving a smaller value than the actual EEG power value out of an arbitrary subset of segments by the law of chance. Under the null-hypothesis, these ranks were assumed to be uniformly distributed between 0 and 1].

The further evaluation steps were adopted from the precedent study (Wackermann et al., 2004). These comprised the transformation of ranks into z-values, the collapsing of data over subjects, the collapsing of data over channels using a chi-square statistic, and the overall assessment of significance for the study. Two different evaluations were performed, based on different time windows: The EEG power was calculated for two different time windows: a) for whole stimulation epochs of one second and b) for a time window of 136ms around the stimulated subject's maximum EEG response (which had been determined individually for each stimulated subject and for each channel electrode position).

For the present analysis, the statistical method was reconsidered in detail, and certain modifications appeared necessary. These are described in the following paragraph.

Modification 1: Sampling for the bootstrap

In what follows, the previous sampling procedure turns out to systematically overestimate effects. From a theoretical point of view, this is due to the fact that the method compared one sample ('stimulation' epochs) drawn out of the data population with a second sample ('interstimulus' epochs) drawn out of the data population by bootstrapping. This violates the premise of exchangeability of epochs between sample and reference (Edgington, 1987). As a consequence of this violation, the second sample, which was used as a reference for the first, does not adequately reflect the possible variability of the first but underestimates it. Therefore, the rank of the first mean within the distribution of means of all subsets is not equally distributed within the interval [0,1] but tends towards the extremes of the interval. This inevitably leads to an overestimation of significance of any observed deviations from randomness in either direction. A simplifying model simulation which uses numeric surrogate data instead of EEG epochs and which calculates *t*-tests for a control of the bootstrap, can illustrate this phenomenon:

A pool of 1,000,000 standard-Gaussian distributed random real numbers served as data population. Out of this pool, a first random sample (sample 1) of $N_1 = 100$ numbers was drawn, for which the mean μ_1 was computed. This was to represent the pool of stimulation period values, but instead of sets of whole EEG segments (for later averaging and calculation of EEG power) only one single value μ_1 for the mean of the drawn numbers was obtained.

Then, a second random sample (sample 2) with $N_2 = 300$ numbers was drawn out of the data population, representing the non-stimulation segments pool (which is larger, about three to one in relation, than the stimulation segments pool). Out of this second sample, 10,000 subsets of numbers were drawn, each as large as sample 1, thus comprising $N_1 = 100$ numbers each. For each subset, the mean was computed, resulting in 10,000 mean values. The cumulative distribution function of these mean values allowed determining the rank R of μ_1 (mean of sample 1) within all of the 10,000 subset means.

According to the pre-assumptions of the applied evaluation method, the rank R should become any value between 0 and 1 with the same probability (even independently from the distribution of values within the data population). After transformation of each rank R into a z-value by the inverse Gaussian function, z-values of multiple repetitions should be distributed standard-Gaussian.

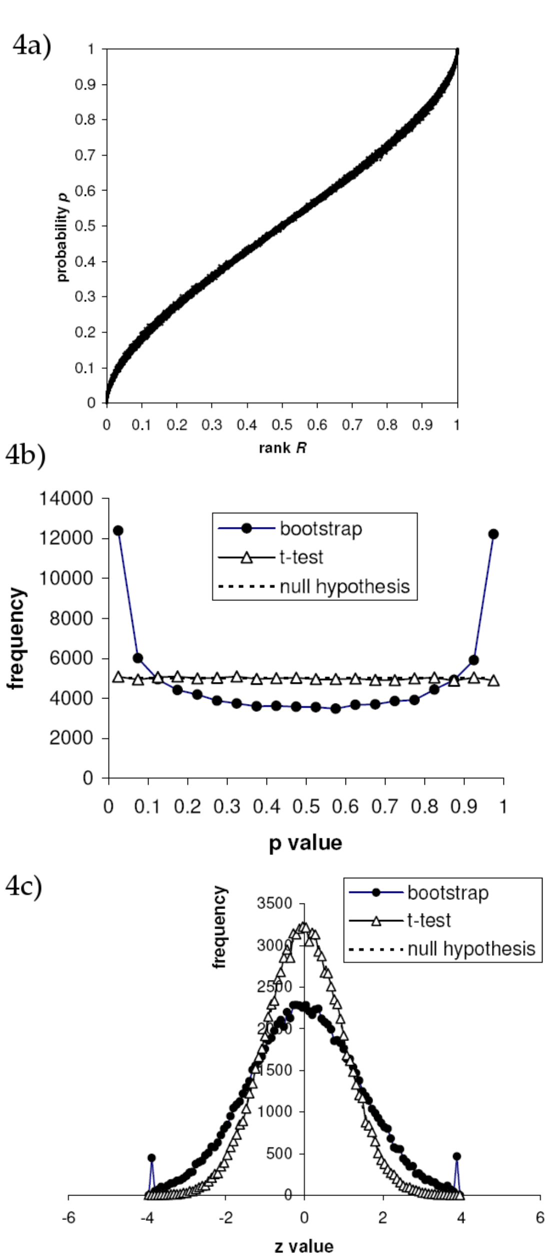
In order to test this, the procedure of drawing samples as described above and calculating a rank R for μ_1 was done repeatedly

100,000 times. Fulfillment of the null hypothesis, meaning that sample 1 and sample 2 were drawn out of the same data population, was guaranteed by the model.

In parallel to the bootstrap-like method, the first (N_1 = 100) and the second random sample (N_2 = 300) were compared for mean difference by using a *t*-test as a conventional reference method. Performing a *t*-test was regarded as permissible, even with different sizes of sample 1 and sample 2, since the data were distributed Gaussian.

Every correct method of testing for mean difference should yield the same value *p*, which is the probability that the first sample is drawn out of a data population with a smaller mean value than the second sample. Given the correctness of the pre-assumptions, the results of the bootstrap and the *t*-test should be identical. The results are depicted in Figure 4. First, the results obtained by both methods (ranks *R* obtained from the bootstrap; *p*-values obtained by the *t*-test) are compared in a scatterplot (Figure 4a); second, the distribution of these values in the interval [0;1] (pre-assumed to be uniform) is plotted (Figure 4b), and third, the distribution of the *z*-transformed values (pre-assumed to be distributed standard-Gaussian) is plotted (Figure 4c). For the bootstrap, the pre-assumption is clearly violated, while it is obviously fulfilled for the *t*-test. Distribution plots for the bootstrap, but not for the *t*-test, show an increased standard deviation (which, according to pre-assumptions, should be 1.0) and a tendency of results towards the extremes.

To rectify the observed bias, the sampling procedure has been modified: the rank *R* of the stimulation epochs EEG power is now calculated within multiple subsets drawn from a reference pool comprising <u>all</u> stimulation and interstimulus epochs of the subject. This modification accomplishes that the reference pool, now comprising all epochs, adequately represents the possible variability of EEG measures obtained from the stimulation epochs. In terms of preassumptions for the bootstrapping procedure, segments are now exchangeable. The effect of this modification has been tested with the aforementioned numeric surrogate data. The results are depicted in Figure 5 (with an arrangement corresponding to Figure 4). The ranks *R* obtained from the bootstrap were now found to be scattered within a small range around the *p*-value obtained from the *t*-test. (Additional



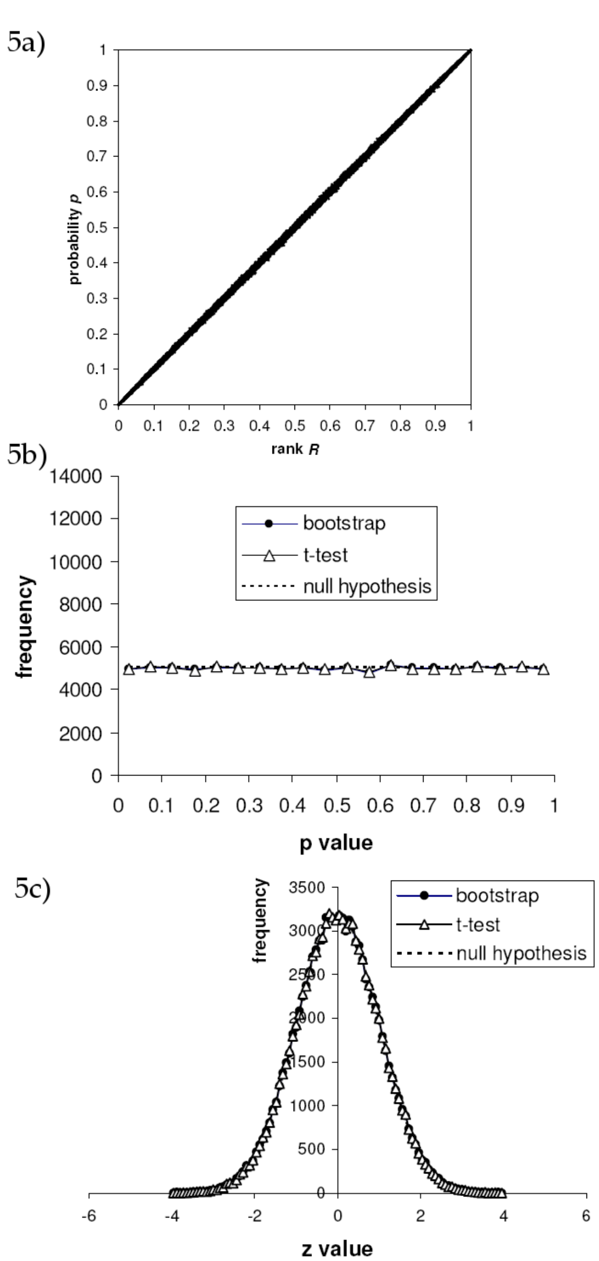


Figure 4. Original method of the bootstrap vs. *t*-test: comparison of results from the numeric simulation. (a) Scatterplot of ranks R (from the bootstrap) versus *p*-values (from the *t*-test). (b) Distribution of the 100,000 obtained ranks R and the corresponding *p*-values yielded by the *t*-test. (c) Distribution of the ranks R (from the bootstrap) versus *p*-values (from the *t*-test) after *z*-transformation, compared with the standard Gaussian distribution. (Note: the two spikes around 4.0 and -4.0 represent cases in which the mean of sample 1 was outside the range of all subset means; instead of using infinite values; they were summarized at the edge of the plotted range for illustration)

Figure 5. Modified method of the bootstrap vs. t-test: Comparison of results from the numeric simulation. (a) Scatterplot of ranks *R* (from the bootstrap) versus *p*-values (from the *t*-test). (b) Distribution of the 100,000 obtained ranks *R* and the corresponding *p*-values yielded by the *t*-test. (c) Distribution of the ranks *R* (from the bootstrap) versus *p*-values (from the *t*-test) after *z*-transformation, compared with the standard Gaussian distribution

analyses revealed that with an increased number of subsets this range converged to zero.)

Furthermore, in order to rule out a weakness in test power, an 'effect' of a known size (e.g. 0.1 standard deviations) was simulated. For this purpose, samples 1 and 2 were drawn out of two different data populations with the means μ_1 and μ_2. The difference between μ_1 and μ_2 was varied in order to simulate different effect sizes. The results showed that bootstrap and t-test yielded the same p values, which suggests that the modified sampling procedure does not bias results in the case of an existing effect.

Modification 2: Correction for degrees of freedom

A second concern refers to the collapsing of results over EEG channels. The z-values from the nineteen EEG channels are not independent from each other; the columns of z-values, each comprising the results from a single EEG channel for all subjects, are considerably intercorrelated. This implies that the chi-square resulting from these z-values has less than nineteen degrees of freedom. Accordingly, the application of a chi-square test according to the preassumption of nineteen degrees of freedom (which would be appropriate in the case of independent channel results) will overestimate the significance of any observed deviation from randomness. A correction for the degrees of freedom was achieved by using a Monte-Carlo simulation.

In multiple (5000) permutations of all EEG segments of a subject, segments were assigned randomly to either the 'stimulation' or the 'interstimulus' epoch pool; pool sizes N_S and N_I were adopted from the real study data. This assignment was performed once per subject and kept constant over the nineteen EEG channels (in order to adopt the degree of interdependence of the channelwise results from the study data). For each such permutation, data were analyzed as described above, which, after collapsing over subjects and over channels, resulted in a chi-square value for each time window (136 ms, 1000 ms) and each condition (uncovered, covered, and difference). The cumulative distribution function for the 5000 collected chi-square values was plotted. From this curve, for any chi-square value, a corresponding p value can be determined, which indicates the probability of not exceeding the according chi-square value by randomness.

For the chi-square values obtained as study results, the corresponding *p* values were readable from the cumulative distribution functions that resulted from the simulation. This allowed an assessment of significance which was corrected for the degrees of freedom of the EEG channels.

Results

All analyses have been performed for two different time windows, 136 ms and 1000 ms; the results are described for both 'uncovered' and 'covered' conditions, as well as for the difference between conditions. Channelwise results after collapsing *z*-values over subjects are summarized in Figure 6. Chi-square values, resulting from collapsing data over channels, are included for each time window and condition.

136 ms

uncovered

	Fp1		Fp2	
	-1.046		-2.008	
F7	F3	Fz	F4	F8
-0.862	0.282	-1.447	-1.697	-0.136
T7	C3	Cz	C4	T8
1.681	0.878	0.107	-0.474	0.190
P7	P3	Pz	P4	P8
0.261	0.209	0.083	0.953	-0.461
	O1		O2	
	-0.984		-0.171	

$\chi^2 = 17.05$

covered

	Fp1		Fp2	
	0.153		1.025	
F7	F3	Fz	F4	F8
0.793	-0.027	0.705	1.711	-0.543
T7	C3	Cz	C4	T8
-0.790	-0.828	-0.900	-1.173	0.401
P7	P3	Pz	P4	P8
-0.978	-1.280	0.087	0.163	-0.180
	O1		O2	
	0.089		0.439	

$\chi^2 = 11.94$

difference

	Fp1		Fp2	
	-0.848		-2.145	
F7	F3	Fz	F4	F8
-1.170	0.218	-1.522	-2.410	0.288
T7	C3	Cz	C4	T8
1.747	1.206	0.712	0.494	-0.149
P7	P3	Pz	P4	P8
0.876	1.053	-0.002	0.558	-0.198
	O1		O2	
	-0.758		-0.432	

$\chi^2 = 23.21$

1000 ms

uncovered

	Fp1		Fp2	
	-0.183		-0.250	
F7	F3	Fz	F4	F8
-1.262	0.995	0.398	-0.319	-1.521
T7	C3	Cz	C4	T8
0.551	1.522	1.830	1.111	-1.004
P7	P3	Pz	P4	P8
0.988	1.159	1.730	1.818	0.692
	O1		O2	
	0.076		0.865	

$\chi^2 = 23.31$

covered

	Fp1		Fp2	
	-0.153		1.545	
F7	F3	Fz	F4	F8
-0.014	-0.954	0.518	0.145	0.101
T7	C3	Cz	C4	T8
-1.163	-1.915	-0.954	0.083	-0.451
P7	P3	Pz	P4	P8
-2.544	-1.727	-1.580	-0.895	-0.109
	O1		O2	
	-2.328		-0.194	

$\chi^2 = 27.98$

difference

	Fp1		Fp2	
	-0.021		-1.269	
F7	F3	Fz	F4	F8
-0.882	1.378	-0.085	-0.328	-1.147
T7	C3	Cz	C4	T8
1.212	2.431	1.969	0.727	-0.391
P7	P3	Pz	P4	P8
2.497	2.041	2.341	1.918	0.567
	O1		O2	
	1.700		0.749	

$\chi^2 = 40.99$

Figure 6. Results after collapsing over subjects: Channelwise z-values for the 136 ms and the 1000 ms time window: 'uncovered' and 'covered' condition, and 'difference between conditions'. Chi-square results after collapsing results over channels are included.

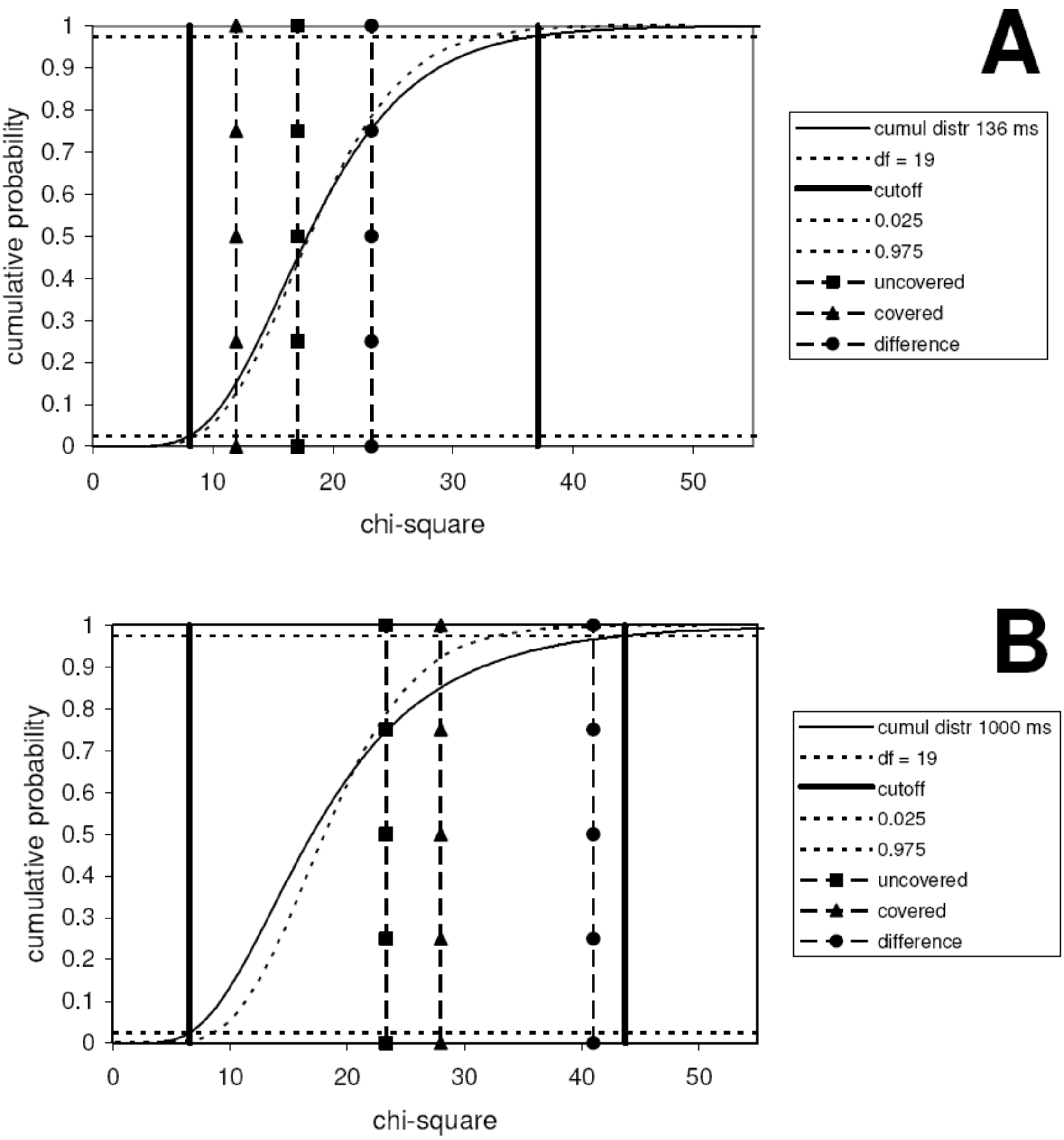


Figure 7. Cumulative chi-square distributions from the Monte-Carlo simulations; A: 136 ms window, B: 1000 ms window. Vertical, dashed lines denote the chi-square values obtained from the study for uncovered and covered conditions and the difference between them. Corresponding *p* values for any chi-squares are read from these distributions; vertical, solid lines mark the lower and upper cutoff chi-squares according to the significance level of 0.05 (two-tailed test). The dotted curve reflects the cumulative chi-square distribution for nineteen degrees of freedom (reflecting the evaluation without correction for degrees of freedom)

Monte-Carlo simulations for both time windows were performed in order to estimate the statistical significance of the chi-squares (i.e. the probability with which pure randomness would produce larger effects than the study data). The results from the Monte-Carlo simulations serve as a reference for the chi-squares obtained from the study data, as depicted in Figure 7. It can also be read from the figure, which excessive significance levels would correspond to the obtained chi-squares without correction for the degrees of freedom (dotted lines).

Comparison with the previous method

In order to estimate the effect of the methodological modifications on the study results, the study data were also analyzed using a) the previous method without modifications ('original method'), b) the previous method with change of only the sampling for the bootstrap, but without correction for degrees of freedom ('new sampling'), and c) the modified method ('new method'). Chi-squares and significance levels obtained from a), b), and c) are compared in Table 1. Note that 'new sampling' and 'new method' do not differ with respect to chi squares but with respect to significance levels.

Table 1: Dependence of study results on the evaluation method: Chi-squares and significance levels (*p* values for two-tailed testing, in brackets) obtained from the original method, obtained after the modification of only the sampling procedure ('new sampling'), and obtained after both modifications of the method ('new method')

	136 ms			1000 ms		
	Original method	New sampling	New method	Original method	New sampling	New method
Uncovered (u)	21.43	17.05	17.05	33.91	23.31	23.31
	(.627)	(.827)	(.898)	(.038)	(.448)	(.511)
Covered (c)	15.45	11.94	11.94	29.00	27.98	27.98
	(.623)	(.224)	(.303)	(.132)	(.168)	(.297)
Difference (u-c)	30.69	23.21	23.21	49.66	40.99	40.99
	(.087)	(.456)	(.492)	(.0003)	(.0048)	(.067)

A further Monte-Carlo simulation was performed to investigate the proportion of positive study outcomes with the previous method under fulfillment of the null-hypothesis. The latter was achieved by multiple random permutations of all 'stimulation' and 'interstimulus' EEG epochs of a subject. Strikingly, the evaluation using the previous method without modification falsely led to 20.1% (26.2%) significant results per condition for the 136 ms (1000 ms) time window, when the significance level α was actually set to 0.05.

Discussion

The present study on correlations in the EEG of spatially separated subjects aimed at a replication and possible clarification of the effects described by Wackermann et al. (2004). The research paradigm was adopted from there.

The EEG power of the non-stimulated subject did not significantly differ between stimulation and interstimulus periods of the stimulated subject. This overall result was obtained after two modifications of the statistical method: A change in the sampling procedure for the bootstrapping, and a correction of the chi-square test for the degrees of freedom of the nineteen EEG channels.

It has to be noted that the chi-square for the 1000 ms time window and 'difference between conditions' just missed the significance level and therefore might remain an object of discussion. On the other hand, a correction of the significance levels for multiple testing has to be considered, because a number of separate evaluations and significance assessments were performed a) for two different time windows and b) for the 'uncovered' and the 'covered' condition and for the difference between conditions. However, even without a correction for multiple testing, none of the six chi-squares exceeded the significance level of 0.05.

The overall result is seen as a negative outcome of the attempt to replicate prior findings of EEG correlations in this specific experimental paradigm. Of course, it may not be interpreted as a proof of the absence of the phenomenon under question; firstly and trivially, the absence of an effect cannot be proven by an unsuccessful replication attempt; secondly, the outcome refers to the specific constellation of EEG correlations if one subject is visually stimulated by checkerboard reversal, which is an abstract pattern.

It is further suggested that the results presented here should not be discussed in terms of 'telepathy', which was the motivational origin of some early studies on EEG correlations between separated subjects (e.g. Grinberg-Zylberbaum et al., 1994). The recent discussion in the field rather refers to 'non-local correlations between distant brains' which is a broader concept than the assumption of a directed transmission of information by means of telepathy. Regardless of the viewpoint however, the present study does not contribute to substantiate the assumption of unexplainable non-local correlations.

Besides the failed replication, the study critically revised the statistical procedure applied in previous studies. It revealed the necessity to introduce two important changes.

With the previous method being applied without modification, results for the 1000 ms (whole-epochs) evaluation suggested a significant change in the non-stimulated subject's EEG power in the 'uncovered' condition, with a highly significant difference between 'uncovered' and 'covered' conditions. The change consisted of an increase in EEG power in parietal, predominantly left-hemispheric regions and a decrease in bilateral temporal regions, which is quite different from the topography of the effects reported by Wackermann (2004). However, using the modified analysis, all observed changes in EEG power were below significance level. Therefore, the results obtained from the previous method are likely due to inadequate statistical analyses.

It is expected that the results of those preceding studies, which used the statistical method discussed here without the two proposed modifications, and which claimed effects unexplainable by conventional physiology and physics, will undergo a decrease in effect size and significance when the revised method is applied on these data. From here, it cannot be estimated with precision, to what extent the previous results have been biased by the statistical method, and to what extent either of the modifications would affect these previous results. However, both of the methodological annotations must have led into the same direction, namely to overestimate the effects under investigation.

References

Charman, R.A. (2006). Has direct brain to brain communication been demonstrated by electroencephalographic monitoring of paired or group subjects? *The Journal of the Society for Psychical Research* 70, 1-24.

Edgington, E.S. (1987). *Randomization Tests*. New York, NY: Marcel Dekker Inc.

Fenwick, P., Vigus, N., and Sanders, S. (1998). *The transferred potential*. Unpublished manuscript, London.

Grinberg-Zylberbaum, J., Delaflor, M., Attie, L., and Goswami, A. (1994). The Einstein-Podolsky-Rosen Paradox in the brain: the transferred potential. *Phys Essays*, 7, 422-427.

Jasper, J.J. (1958). The 10/20 international electrode system. *Electroencephalogr. Clin. Neurophysiol.*, 10, 371-375.

May, E. C., Spottiswoode, S. J. P., and Faith, L. V. (2001). A methodological issue in the study of correlation between psychophysiological variables. *Proceedings of the 44th Annual Convention of the Parapsychological Association*, 166-178.

Sabell, A., Clarke, C., and Fenwick, P. (2001). Inter-subject EEG correlations at a distance - the transferred potential. *Proceedings of the 44th Annual Convention of the Parapsychological Association*, 419-422.

Standish, L. J., Kozak, L., Johnson, L. C., and Richards, T. (2004). Electroencephalographic evidence of correlated event-related signals between the brains of spatially and sensory isolated human subjects. The *Journal of Alternative and Complementary Medicine*, 10, 307-314.

Radin, D.I. (2003). Thinking outside the box: EEG correlations between isolated human subjects. *Proceedings of the 43th Annual Convention of the Parapsychological Association*, 184-199.

Radin, D. I., (2004). Event-related electroencephalographic correlations between isolated human subjects. *The Journal of Alternative and Complementary Medicine*, 10, 315-323.

Wackermann, J., Seiter, C., Keibel, H., and Walach, H. (2003). Correlations between brain electrical activities of two spatially separated human subjects. *Neuroscience Letters*, 336, 60-64.

Wackermann, J. (2004). Dyadic Correlations between Brain Functional States: Present Facts and Future Perspectives. *Mind and Matter*, 2, 105-122.

Wackermann, J., Naranjo Muradás, J.R., and Pütz, P. (2004). Event-Related Correlations between Brain Electrical Acitivites of Separated Human Subjects: Preleminary Results of a Replication Study. *Proceedings of the 47th Annual Convention of the Parapsychological Association*, 465-468.

European Journal of Parapsychology
Volume 23.2, pages 147–153

ISSN: 0168-7263

Dyadic EEG correlations re-examined: A commentary on the replication study by W. Ambach

Jiří Wackermann
Department of Empirical and Analytical Psychophysics
Institute for Frontier Areas of Psychology and Mental Health
Freiburg i. Br., Germany

Some years ago we were reporting results of an electrophysiological study (Wackermann et al., 2003) suggesting a relation between brain states of two separated subjects, assessed by parallel EEG measurements from both subjects. In an attempt to replicate the finding with an improved experimental setup (Wackermann et al., 2003), we again found deviations of our measure of interest from the null-hypothesis expectation, but showing a confusing (if any) pattern difficult to reconcile with our earlier findings. We were facing a situation indicating either a 'paranormal' nature of the alleged effect, or its possibly artefactual origin. A follow-up study in our laboratory could not be carried out as planned due to personnel changes in our department. Therefore, the opportunity to arrange a replication study in co-operation with the research group for Clinical and Physiological Psychology was welcomed. True replication studies are rather rare; now there was an excellent opportunity to test whether the same (or at least qualitatively similar) results can be obtained in an independent study, in another laboratory and with a different equipment, but following an identical experimental protocol. The results of the replication study conducted by Dr Ambach are presented as a separate paper (Ambach, 2008), and we shall just add a few comments.

Address correspondence to: Jiří Wackermann, Institut für Grenzgebiete der Psychologie und Psychohygiene, Wilhelmstr. 3a, D-79098 Freiburg, Germany. E-mail: jw@igpp.de .

Firstly, some introductory remarks for a reader who may not be familiar with the research topic and its history. The problem under study was introduced by Grinberg-Zylberbaum & Ramos (1987) and Grinberg-Zylberbaum et al. (1994) who were searching for physiological indicators of hypothetical non-sensory, 'direct communication' between two subjects. The term 'correlation' used in the title of the commented paper and our earlier communications does not refer to cross-correlations between native EEG signals. Rather, it refers to correlation between *brain states* (see Wackermann, 2004), induced by sensory stimulation at one of the subjects (A), and possibly revealed by a suitably chosen signal parameter in the brain electrical activity of the other, non-stimulated subject (B). Stimulus-onset aligned data averaging is a standard technique used in studies of stimulus-related brain electrical responses (also called 'evoked potentials', EP). The experimental paradigm used in our studies was derived from that used by Grinberg-Zylberbaum et al. (1994), who reported an EP-like activity in the subject B's averaged EEG recorded synchronously with subject A's stimulation. We could not confirm Grinberg-Zylberbaum's finding of a 'transferred potential', but we observed subtle power fluctuations of the averaged EEG signal (Wackermann et al., 2003).[1]

Specifically, the parameter of interest was the effective voltage ratio, $Q \equiv V_{\mathrm{eff}}/V_{\mathrm{ref}}$, where V_{eff} was the effective voltage determined for the averaged N_{S} EEG data segments recorded during the visual stimulation of the other subject (briefly: stimulation segments), and V_{ref} was a reference value determined from the same subject's EEG data recorded during the inter-stimulation periods. The data-analytic strategy applied in our studies relied upon an elementary theorem of probability theory (Feller, 1950): Let X be a continuous random variable with a cumulative distribution function (CDF) $F_X(x) \equiv \mathsf{P}(X < x)$; then the transformed variable $U = F_X(X)$ is uniformly distributed on the interval [0,1], so the CDF of U is $F_U(u) \equiv \mathsf{P}(U < u) = u$; a principle underlying all 'tests of significance'.[2] Since we had no prior knowledge or a parametric model

[1]The exclusive focus on averaged EEG data was occasionally and deservedly criticised. It should be clear from the above-said that this choice was historically determined: if we were designing a data-analytic approach from scratch, certainly we would not begin with data averaging. This observation also reveals limitations of the empiricist approach, building experiments by variations of others' experiments rather than on prior theoretical reasoning.

[2]Where usually the CDF of a test statistics, say, T, under a given hypothesis is derived analytically and used to transform the observed values of T to $U = F_T(T)$ or, by convention, to its complement $P = 1 - U$. It is the uniformity of the CDF of U what allows us to read the P-values directly as 'probabilities' of reaching the observed value of statistics T.

of the distribution of Q, *empirical* CDFs of Q were constructed from repeated random drawings of inter-stimulation EEG data sub-sets from a pool of a total size $N_I > N_S$. These CDFs were used to transform the observed Q values for the stimulation EEG data, *i.e.*, to determine their ranks R with respect to a 'reference population' of non-stimulation Qs. [The procedure was summarised in (Wackermann et al., 2004), Fig. 1, reproduced in the commented paper as Fig. 3.] For the purpose of final statistics, it was assumed that, under the null hypothesis, the ranks $R = F_Q(Q)$ are *uniformly* distributed on [0,1]. Under this assumption, values $Z = \Phi^{-1}(R)$ would be *normally* distributed.

Ambach's numerical simulations, based on random drawings from a Gaussian pool (pp. 8*–9*) show that this assumption not necessarily holds: the distribution of ranks evaluated with respect to empirical CDFs may be non-uniform (p. 10*, Figs. 4a).[3] The real distribution of Rs shows a relative excess at the lower and upper ends of the [0,1] interval, which results in a ∪-shaped PDF reminding of arcsin distribution. The deviation from uniformity as such is not surprising; it can be easily predicted for small samples from the discrete character of the ranks. What is rather counter-intuitive — at least for a worker used to rely upon asymptotic statistics — is that with increasing size of both samples the curvature does not flatten to a uniform distribution; it only becomes smoother but preserves its shape. This argument is essentially *qualitative*: it says that the deviation from uniformity may *in principle* bias the rank-based statistics; the outcome of the simulation is presented 'as is', and no rationale is given for the particular choice of simulation parameters.

It is, however, instructive to examine the distributions of empirically determined ranks with varied sample sizes. We used the same simulation procedure as Ambach, with only slightly different sum $N_1 + N_2$ = 360, instead of 400, and partitioned the total 360 in varied ratios. The simulations[4] (Fig. 1) reveal that the magnitude of deviation from uniformity depends essentially on the samples size ratio N_1/N_2. For ratio 1/3 (Fig. 1, upper row, middle), the results agree with those reported by Ambach; the deviation is even more pronounced for ratio 1/2 (upper row, middle). With decreasing N_1/N_2 ratio, however, the CDF comes closer to the ideal form (see Fig. 1, lower row), so the deviation of Z-transformed

[3] In our view, the co-lateral argument based on parallel t-tests is rather diverting and obscuring the central message, i.e., the deviation of ranks distribution from uniformity. We do not need a t-test to know the theoretically correct distribution.

[4] The source C code of the simulation program is available on request from the author.

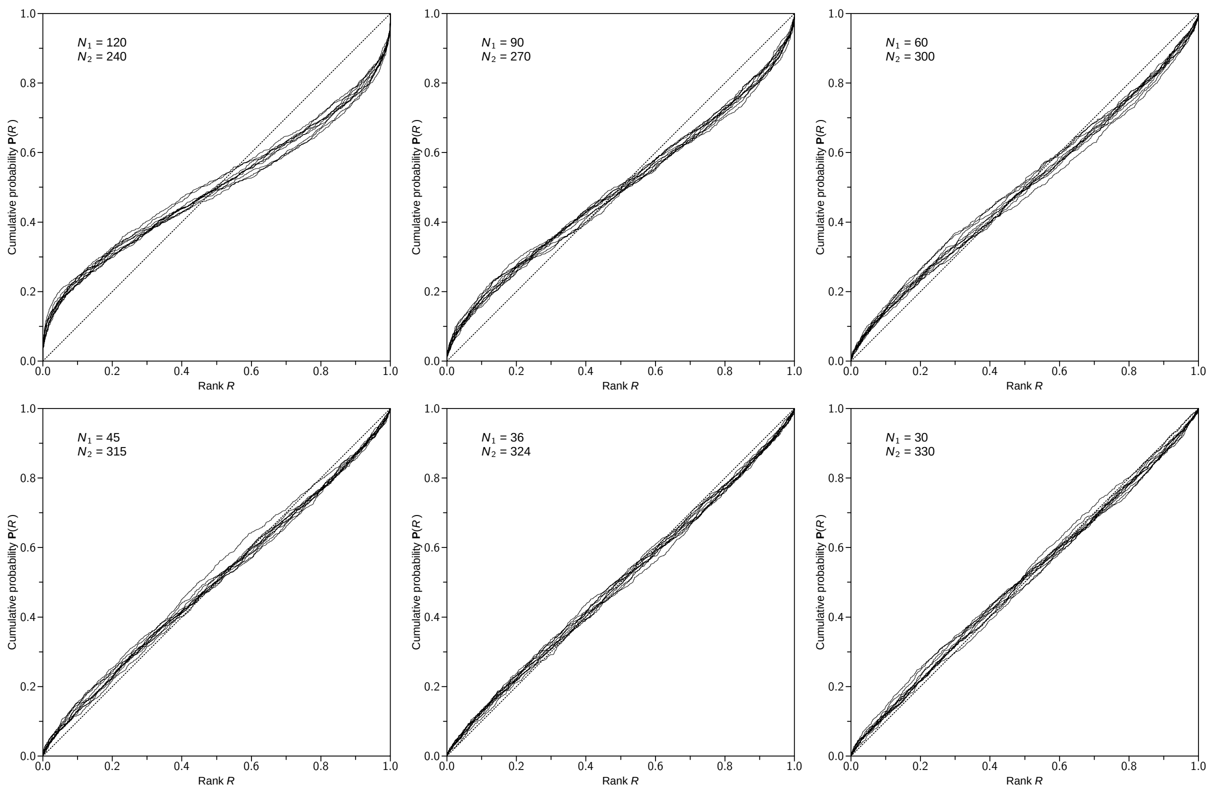


Figure 1. Cumulative distribution functions of ranks R determined by numerical simulations according to Ambach's procedure, with varied sample sizes N_1 ('experimental' sample) and N_2 ('reference' sample). Ratios N_1/N_2 are (upper row) 1/2, 1/3, 1/5, (lower row) 1/7, 1/9, and 1/11. The dashed diagonal lines indicate the ideal (uniform) CDF. Note that from $N_1/N_2 \approx 0.1$ on the deviation is almost negligible.

ranks, $Z = \Phi^{-1}(R)$, would diminish with $N_1 \ll N_2$. It follows, then, that the rank-based procedure used in our earlier studies is in principle sound *if the samples size ratio is taken into account*. The ratio N_S/N_I (stimulation *vs.* inter-stimulation data segments) can be roughly determined by the experimental design.[5] Alternatively, the deviation from the uniform distribution could be modelled parametrically and corrected for in subsequent steps of the analysis.

At any rate, Ambach's study has pointed out that more attention has to be paid to the construction of data-analytic methods used in experimental studies of psycho(physio)logical or biophysical 'anomalies'. It has demonstrated that apparently 'intuitive' statistical properties must not be assumed without careful verification, especially where relatively small samples ($N \approx 10^1$ — 10^2) are concerned: an important aspect that should be appreciated. However, the insights derived from

[5]We say 'roughly' because the exact value will depend on the data loss rate due to artifact elimination, and is thus not perfectly under experimenter's control.

the simulations should not be applied too hastily. In the commented paper, the bare fact that P-values obtained from empirical distribution functions by our original method *may be biased* was turned into an argument for the *necessity* of merging the two data subsets (stimulation and inter-stimulation EEG segments), and for using the combined dataset as a reference pool (p. 11*) — a point on which we disagree. It is difficult to resist an impression that the eventually successful 'deconstruction' of the effect under study serves here as a proof of statistical reconsiderations and modifications based thereupon. We would prefer a systematic study of the bias, focussing particularly on (i) a quantification of the bias as to its dependence on the dataset size, and (ii) a more convincing demonstration that the model used in demonstration of the bias applies to analyses of real EEG data.

For a reader who may find our assessment slightly opinionated we wish to add that we do not argue in favour of the 'reality' of the effect under question. In fact, from our operationalist stance the phrase 'really existing effect' has no other meaning that 'consistently reproducible phenomenon' (given that the necessary conditions for the phenomenon are known). Reviewing results of our own replication study (Wackermann et al., 2004), and taking into account the results reported by Ambach *even without post hoc corrections*, it is highly doubtful that there is anything such as a 'real' effect. This negative evidence arises from the remarkable lack of consistency, in terms of direction and spatial distribution of the effect measures, and is thus based more on a visceral feeling for 'what is like real physiology' than on formal statistical inference.

Our final comment is not related directly to the commented study but, generally, to studies of correlations between physiological activities of separated subjects, or any biophysical 'anomalies' of similar sort. During the last few decades, the search for so-called 'physiological signs of psi' (Beloff, 1974) went through several waves of enthusiasm, often elicited by novel experimental approaches ('direct communication', 'direct mental interaction between living systems', 'presentiment', etc), and sometimes claimed to promise a fundamental revision of our view of nature (e.g., quantum-like 'entanglement' between human brains, 'reversal of time arrow', etc).

As to our knowledge, none of those high hopes has ever been fullfilled, and none of those approaches developed into a really working experimental paradigm — that is, one yielding reproducible results across laboratories, results that would visibly stand out of the bush of error bars. There are no signs of real progress. May each individual researcher draw her/his conclusions from this observation. We take the lesson seriously and turn to more productive research topics, not to spend our lives in a heroically 'relentless' but ultimately unproductive search.

Acknowledgements

Thanks to Dr Ambach for intense discussions of the topical issues, and to two reviewers for helpful remarks on this commentary.

Abbreviations and mathematical symbols

CDF	cumulative distribution function
EEG	electroencephalogram
EP	evoked potential
PDF	probability density function
F_X	CDF of random variable X
Φ	CDF of normally distributed random variable
Φ^{-1}	inverse function to Φ
N	sample size
P	probability of an observed event
P	probability associated with a test statistics
Q	effective voltage ratio
R	rank value
U	uniformly distributed random variable
V_{eff}	effective voltage (stimulation)
V_{ref}	reference effective voltage (inter-stimulation)
X	random variable, general
Z	random variable, normally distributed

References

Ambach, W. (2008). Correlations between the EEGs of two spatially separated subjects — A replication study. *European Journal of Parapsychology*, this issue.

Beloff, J. (1974). ESP: The search for a physiological index. *Journal of the Society for Psychical Research*, *47*, 403–420.

Feller, W. (1950). *An Introduction to Probability Theory and Its Applications*, Vol. I. New York: Wiley.

Grinberg-Zylberbaum, J. & Ramos, J. (1987). Patterns of interhemispheric correlation during human communication. *International Journal of Neuroscience*, *36*, 41–53.

Grinberg-Zylberbaum, J., Delaflor, M., Attie, L., & Goswami, A. (1994). The Einstein-Podolsky-Rosen paradox in the brain: The transferred potential. *Physics Essays 7*, 422–427.

Wackermann, J., Seiter, C., Keibel, H., & Walach, H. (2003). Correlations between brain electrical activities of two spatially separated human subjects. *Neuroscience Letters*, *336*, 60–64.

Wackermann, J. (2004). Dyadic correlations between brain functional states: Present facts and future perspectives. *Mind and Matter*, *2*(1), 105–122.

Wackermann, J., Naranjo Muradás, J. R., & Pütz, P. (2004). Event-related correlations between brain electrical activities of separated human subjects: Preliminary results of a replication study. In S. Schmidt (Ed.), *Proceedings of the 47th Annual Convention*, Vienna: Parapsychological Association, pp. 465–468.

European Journal of Parapsychology
Volume 23.2, pages 154-172

ISSN: 0168-7263

Lunar-magnetotail Encounters as Modulators of Mind-Matter Interaction Effects

Eckhard Etzold
Society for Anomalistics, Postfach 1202, D-69200, Sandhausen, Deutschland

Abstract

Mind-matter interaction (MMI) effects are controversial partially due to a lack of a theory that persuasively explains such effects, and also due to the difficulty of replicating empirical effects in controlled experiments. One potential explanation for this empirical capriciousness might be the presence of physical factors which modulate MMI performance. One previously suggested variable has been the Earth's geomagnetic field; another is the lunar phase. The hypothesis that the Moon's interaction with the Earth's magnetosphere modulates MMI performance was tested in data collected in a long-term, online MMI experiment. The analysis showed a clear influence of the Earth's magnetotail in the MMI results, confirming the hypothesis. This suggests that phenomena relying on purported MMI may be more efficacious during quiet geomagnetic periods of the full moon, when the Moon is passing deep through the inner plasma sheet of the magnetotail

Introduction

Mind-matter interaction (MMI) is the hypothesized ability of the mind to influence matter or energy without the use of any currently known physical means. Spiritual healing, psychokinesis, distant healing, and unusual human-machine interactions might be based on MMI effects (Heath, 2003). Although hundreds of laboratory studies seem to provide evidence of MMI, the existence of such effects remains controversial because there are no convincing theoretical reasons to explain how such effects may exist (Bösch, Steinkamp, Boller, 2006; Radin, Nelson, Dobyns, Houtkooper, 2006). Barring theory, the central

Correspondence details: Eckhard Etzold, Grosse Grubestr. 2A, D-38122 Braunschweig, Deutschland. Email: eckhard.etzold@gmx.de

issue in this line of research has been the question of reliability. For example, a large-scale, multi-laboratory replication effort failed to repeat previously successful MMI experiments reported by Princeton University’s Engineering Anomalies Research Laboratory (Jahn et al., 2000). While critics might point to the failure to replicate as evidence that MMI does not exist, another interpretation is that we do not fully understand the conditions under which MMI is best demonstrated. This prompted a search for environmental parameters which modulate the MMI performance via enhancing or reducing human powers of concentration.

Geomagnetic effects on human behaviour

The influence of subtle electromagnetic effects on our human nervous system has been a topic of scientific research for decades. Evidence of an effect of ELF electromagnetic fields on human pineal gland function was reported (Wilson et al., 1990) and rejected (Gobba, Bravo, Scaringi & Roccatto, 2006). Persinger found correlations of geomagnetic activity with enhanced anxiety, sleep disturbances, altered moods, and greater incidences of psychiatric admissions (Persinger, 1987). Researchers had reported that poltergeist episodes frequently begin on the day of a sudden and intense increase in global geomagnetic activity (Gearhart & Persinger, 1986). Sturrock (2004) analysed data of UFO events and found significant correlations with local sidereal time. Spottiswoode (1990) reported the existence of a negative correlation between scores in free response anomalous cognition experiments and geomagnetic fluctuations. Sudden infant deaths seem to be associated with continuous micropulsations in times when global geomagnetic activity is very low (O'Connor & Persinger, 1999). Increased solar and geomagnetic activity seems to be associated with increased arterial blood pressure (Ghione, Mezzasalma, Del Seppia & Papi, 1998). Distinct effects of the Moon on geomagnetic activity were also reported (Bigg, 1963; Bell and Defouw, 1964; Knott, 1975).

Despite the large volume of reports, in total we do not really understand the role of these electromagnetic and geomagnetic interactions on human behaviour. Neither do we understand the underlying mechanism of MMI effects. As a result, attempts to explore MMI effects by using patients as targets might raise serious ethical problems, since we don't know when MMI effects might be healing or harming others. All that have today is nothing more than some small pieces of a jigsaw. The use of MMI experiments with a random number generator as a target for the evaluation of the effectiveness of global geomagnetic activity on human behaviour does minimize the risk of adverse effects on the health of subjects and is thus ethically safe. It is a great way to learn more about this still unexplored realm of human ability and behaviour.

Lunar phase effects in MMI data

In an earlier study, Radin (1997) claimed evidence of MMI effects in casino pay-out rates which depended on the lunar phase. The peak effect was found within one day of the full moon. Based on many previous studies examining correlations between the strength of the Earth's geomagnetic field (GMF) flux and purported psychic experiences, it was predicted that the casino pay-out rates would be negatively correlated with GMF. Radin found a higher than average payout rate on full moon days of quiet geomagnetic activity, and a lower than average payout rate on full moon days with high geomagnetic activity. In both cases yet, the correlation would be negative. Thus, we may assume that the effects might cancel each other out.

These results might indicate that the human power of concentration is affected by a geomagnetic parameter. Recently, Sturrock and Spottiswoode (2007) reported a significant lunar-periodic effect in free response anomalous cognition experiment data, which might also confirm this relationship of a lunar-geomagnetic parameter and the human power of concentration.

Attempts to independently check Radin's claims using data from a large-scale, online MMI experiment revealed evidence for a complex solar-periodic full moon effect (Etzold, 2000; Etzold, 2002).

The Fourmilab Retropsychokinesis Project

Etzold (2005) analysed data of the Fourmilab RetroPsychoKinesis Project, an online experiment which "explores the purported anomalous effect known as retropsychokinesis"[1], also known as

[1]http://www.fourmilab.ch/rpkp/proposal.html

"retroPK". RetroPK is a hypothesized ability to retroactively influence random data via MMI. Another hypothesis, the Decision Augmentation Theory (DAT), attempt to explain retropsychokinesis effects with an anomalous information transfer by decision augmentation. This means that people are somehow able to foresee the future and intuitively feel which goal of the experiment might be more promising (May, Utts & Spottiswoode, 1995). If the DAT hypothesis is true, one would expect in case of strong MMI effects in the experimental data a reverse effect in control data taken from the same database. Since DAT can be reduced to intuitive data selection, this might be in fact a kind of data splitting. The lack of any reverse effects in control data might indicate that the hypothesis of retroactive influence is more qualified to explain the effects than intuitive data selection.

Random data for the Fourmilab test are derived from the HotBits hardware random number generator (RNG) based on radioactive decay[2]. Date and time stamps for each experimental result, downloaded from that website, were converted to a lunar phase in degrees. 105 Subsets or bins were created for each lunar cycle from January 1997 to October 2005 (Etzold, 2005). These MMI data were correlated with F10.7 solar radio flux, sunspot numbers, solar wind speed and GMF 'ap'-index data. Significant results were found for MMI data correlated with F10.7 solar radio flux and sunspot numbers. The correlation of the GMF 'ap'-index data with MMI data was barely significant on the $p = .05$-level. This barely significant result might depend on remaining diurnal variations in the GMF 'ap'-index data. Therefore we tried another analysis with reduced diurnal variations in the GMF 'ap'-index data, and we expected an increasing significance for the correlations of the GMF 'ap'-index data with MMI data.

The Moon-magnetotail interaction hypothesis

A hypothesis was suggested by earlier analyses that the Moon's interaction with Earth's magnetosphere during the Moon's passage through the magnetotail in full-Moon times might modulate MMI performance (Etzold, 2005). We wanted to test this hypothesis and see whether high and low MMI effects in the full-Moon interval could be explained by changing geomagnetic activity.

[2] http://www.fourmilab.ch/rpkp/experiments/contents.html

While the effective parameter responsible for modulating MMI performance remains unknown, analyses seemed to indicate that the origin of this parameter may be located in the lunar phase. We assumed more precisely that the actual source of this parameter may be based on lunar-magnetotail interactions.

As illustrated in Figure 1, the Earth's magnetosphere is distorted by the solar wind into a teardrop shape, which extends far into deep space (Tsyganenko, 1995). The magnetotail of this drop stretches in a direction opposite to the Sun, and reaches beyond the lunar orbit. The solar wind follows a path that goes around the Earth in a sort of cover or sheath. This is the magnetosheath plasma region. Between the magnetosphere and the magnetosheath is the magnetopause. The magnetotail is divided into the outer lobes and the inner plasma sheet.

During the full moon interval, the Moon passes the magnetosheath plasma and the lobes of the magnetotail for approximately four days (Schubert, Sonett, Smith, Colburn & Schwartz, 1975).

An indicator for the extent that the Moon spends in the magnetotail is the time the Moon spends in the inner plasma sheath (figure 2). This time varies from 0-15 hours per month at a minimum rising to 60-75 hours per month at the maximum with the 18-year Saros period (Hapgood, 2007) and with the varying solar activity (Tsyganenko and Sitnov, 2007) of the 11-year solar cycle. The lunar orbit is tilted in relation to the ecliptic, and in this way the Moon can pass above, below, or directly through the plasmasheet, depending on the Moon's position in the Saros period.[3] The last period of maximum length of time the Moon spent in the plasmasheet was from 1993 to 2001, and the next period of maximum length is expected around 2011 to 2019. These are very promising time intervals for studying fullmoon-MMI effects. From 2003 to 2010 there is a period of minimum length, leading to the expectation of low fullmoon-MMI effects. But even when the Moon is expected to pass through the magnetotail, a sudden collapse of the entire nightside field with the magnetotail can happen when a strong geomagnetic storm is triggered by high solar activity.

[3] An animated .gif-file created by M. Hapgood, showing the lunar crossing of the magnetotail around the full moon in December, 2007, is available at http://uk.geocities.com/mike.hapgood@btinternet.com/moon_movie_6.gif.

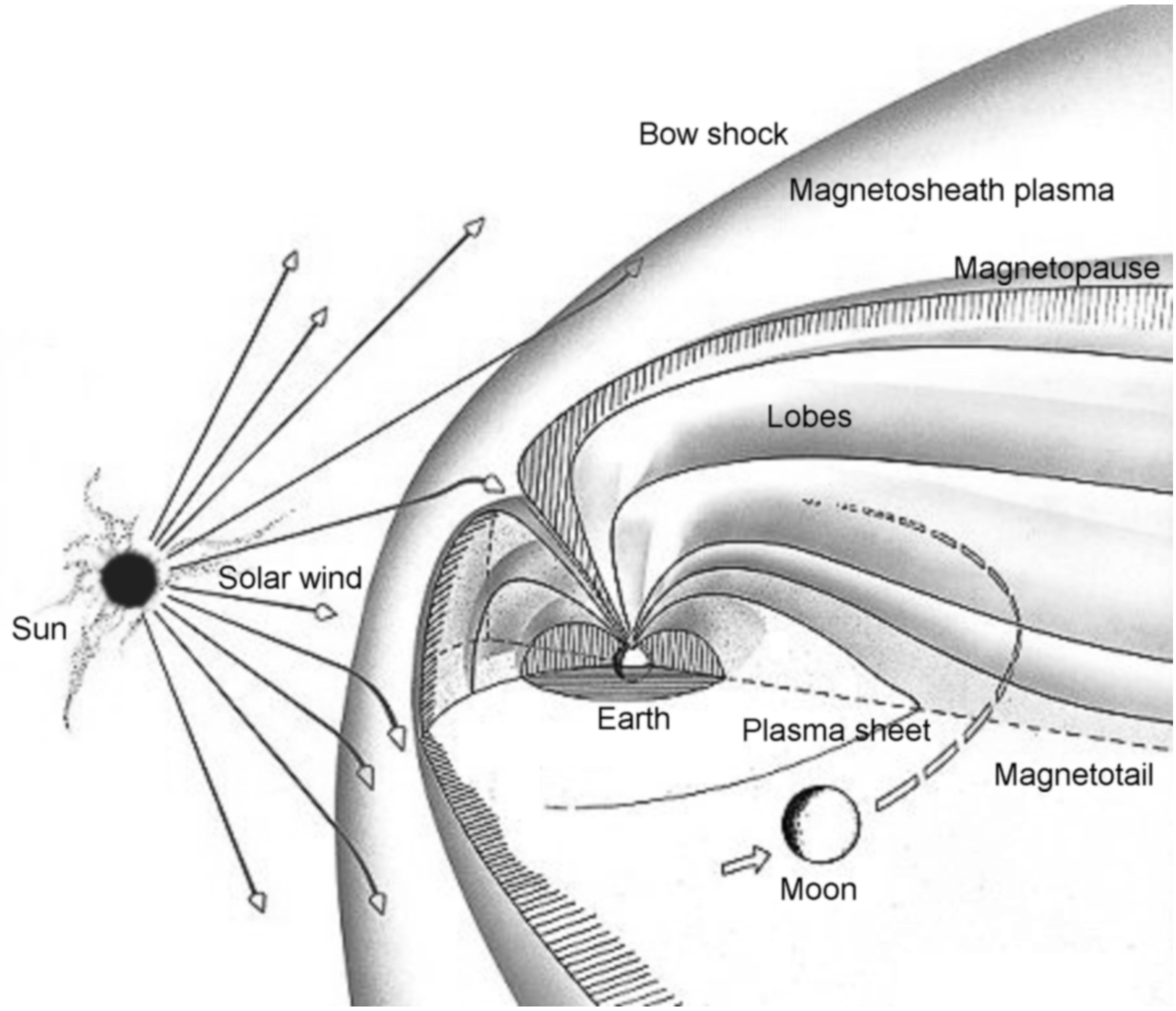


Figure 1. Earth's magnetosphere. Picture was provided by NASA.

In view of the hypothesis that an unknown parameter is modulating MMI performance which results from the Moon's interaction with the magnetotail, one might assume that the correlation of MMI data with GMF 'ap'-index data might yield more significant results. But this was not the case in the Etzold study (2005). Higher significance was found for perihelion parameters like F10.7 solar radio flux and sunspot numbers. The current paper will show that this lack of significance is due to the properties of the GMF 'ap'-index data. Therefore, the data in the present analysis were retrieved from the same retroPK database for the same evaluation period as mentioned before. These are all MMI data from the time period of January 11, 1997, 17:33 UTC through October 8, 2005, 12:45 UTC, all of which fulfil the specifications of the Fourmilab RetroPsychokinesis Project. Newer data were ignored and retained for future replication attempts.

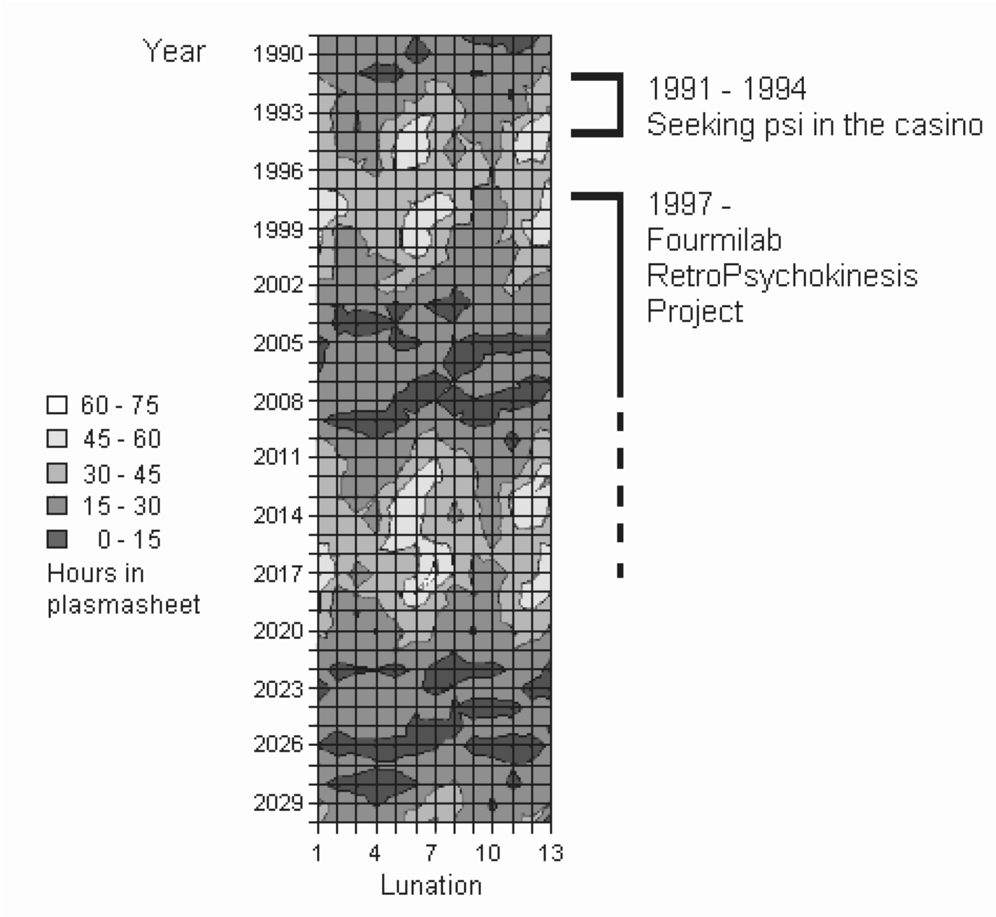


Figure 2. Hours, the Moon spends in the inner plasmasheet. Detail of a plot by Mike Hapgood (2007), and figure was treated by the author. Bright areas mark time intervals of deep lunar encounters with the inner magnetotail, and with a high fullmoon-MMI effect expectation.

Methods

We wanted to see what the solar-periodic full moon effect looks like because the lunar-GMF-MMI hypothesis stated that the Moon's interaction with Earth's magnetosphere during the Moon's passage through the magnetotail at full moon times might modulate MMI performance. We therefore scanned the data with a sliding window function over the whole lunar phase. We used the GMF 'ap'-index data which provide the average global geomagnetic field activity over a three hour period. The GMF 'ap'-index data were obtained in a fixed time interval of every three hours. This results in a constant number of GMF 'ap'-index data for every day of the successive lunar cycles.

Three different data bases are available on the Fourmilab server: MMI 'for the record' data, MMI 'practice' run data and control run data. Here, we are analyzing the MMI 'for the record' data and also 'practice' run data and control run data for a comparison.

The properties of the data

The MMI 'for the record' experiment data were only recorded when somebody was conducting an experiment. Thus, the amount of experimental data varied largely for the same interval of successive lunar cycles with a minimum of 152 and a maximum of 5,555 'for the record' experiments per lunar cycle. In total, 199,632 'for the record' experiments and 202,958 practice run results for the experiment time interval of January 11, 1997 to October 8, 2005 are evaluated. Since December 26, 1997 control experiments have also been running automatically, exactly one run every hour, totalling 59,860 as of October 8, 2005. The first and second generation HotBits configuration used a Windows 95 machine which "crashed every couple of months for no discernible reason"[4]. As a consequence, the Hotbits server was upgraded in September 2006 with a far more reliable Linux-based server (Fedora Core 5).

Further complicating matters is the fact that the control runs do not really match the MMI data. While there are 199,632 MMI experimental data in the evaluation period, only 59,860 control run data are at hand which are a little bit more then a third of the MMI data volume. In addition, practice run results were usually not reported. But due to the important issue of control data, we have to consider every source of information and as a result we looked at the available practice run results.

GMF 'ap'-index data and diurnal effects

The GMF activity varies in diurnal rhythm and with the location on Earth. Besides day and night changes, another reason for diurnal variations is the fact that Earth's geomagnetic poles do not conform with the Earth's geographic poles. As a result, Earth's magnetotail shifts in diurnal rhythms. However, the GMF 'ap'-index data represent a global average of GMF activity as an integer value with minimized

[4]http://www.fourmilab.ch/hotbits/hardware3.html

diurnal rhythm. For example, the exact full moon time might be in one lunar cycle at noon (*e.g.*, 12:00 UTC) with a lower GMF level and in the evening (*e.g.*, 18:00 UTC) of another lunar cycle when the GMF level is higher. Both time intervals might yield the same GMF 'ap'-index data value, but the effect of the lunar/GMF interaction might be completely different. Etzold (2005) computed the correlations of MMI data with GMF 'ap'-index data for every lunar cycle separately. For our evaluation here this means that we have to combine the data of multiple lunar phases in order to minimize diurnal effects and make the data compatible. This means that effects, by accumulation, become evident which are not present if we evaluate data of single lunar phases with much smaller volumes of data. Eight GMF 'ap'-index data values were obtained every day, and reduced diurnal effects in the MMI data can be expected if no fewer than four lunar cycles were included in each subset. This is the case if we divide the entire database into 16 subsets, each with 12,500 MMI 'for the record' experimental data, and the last with 12,132 experimental data. Date and time stamps of these data were converted into lunar phase in degrees and sorted according to ascending lunar phase for each subset.

The control run, practice run, and GMF 'ap'-index data were also divided into 16 subsets but with a different number of experiments per subset with respect to the time intervals of the MMI data to enable comparability of the results. Date and time stamps of these data were also converted into lunar phase in degrees and sorted according to ascending lunar phase for each subset.

Configuring the full moon interval & tests with different sliding window widths

The time interval which the Moon spends in the magnetosphere is estimated at four days. Radin and Rebman (1998) mentioned an interval of one day within the full moon period in which the full moon effect was found in the casino data. Including the day before and the day after the full moon day, this constitutes an interval of three days centered around the time of the full moon. On one day the Moon moves 12.19° along its orbit on average. Thus, Radin's full moon interval is equivalent to 161.7° to 198.3° lunar phase. Etzold (2000) mentioned a lunar phase interval of 166.5° to 192.4° of approximately two days. Since we compute time intervals in lunar phase in degrees, we have 25° lunar phase as a minimum for a two-day interval and 50°

lunar phase as a maximum for a four day interval. For finding the maximum magnitude of an assumed full moon effect in the MMI data, we check all data in steps of 5°, beginning with 25° and ending with 50°. This different window size test should ensure that the effect is not sensitive to the percentage of the lunar cycle that is analysed.

The data of each MMI subset and the corresponding GMF 'ap'-index data subset were sorted according to the the increasing lunar phase with overlapping data for the new moon interval. A computing function with a window width of 25° up to 50°, starting with 0° and up to 360°, overlapping in the start and end-interval, scanned simultaneously the first MMI subset and the corresponding GMF 'ap'-index data subset synchronously from the beginning to the end and calculated the correlation value for every step (0.5°). The sliding window function computed the MMI average effect size and the average GMF 'ap'-index value for every step. For these values, Pearson's r was computed. Since we have 16 subsets, we get 16 Pearson correlation values for each step of the sliding window function. The following checklist displays the basic steps of the analysis:

1. Download the MMI data from the Fourmilab server[1].
2. Convert UNIX time into UTC date and time, and hexadecimal experiment data into a bit score.
3. Divide the dataset into 'for the record' and practice data.
4. Check databases for data which do not conform with the Fourmilab Retropsychokinesis Project specifications.
5. Create 16 subsets with 12,500 MMI experimental data lines each.
6. Convert date and time of each experiment into a lunar phase degree with 0° for new moon, 90° for the first quarter, 180° for full moon and 270° for the last quarter.
7. Sort data according to increasing lunar phase for each subset, starting with 0° and ending with 360°
8. Copy data with lunar phase < 60° and add 360 for every subset. Insert these data at the end of each subset for overlapping scans at new moon position.
9. Perform the same procedure for control run and practice run data.
10. Download GMF 'ap'-index data[5].

[5]The GMF 'ap'-index data were retrieved from the World Data Center for Geomagnetism, Kyoto. http://swdcdb.kugi.kyoto-u.ac.jp/

11. Eight GMF 'ap'-index data were obtained per day. Complete hour and minute information of the middle of each three-hour 'ap'-index interval.
12. Take date and time of the beginning and the end of each MMI data and divide the GMF 'ap'-index data into corresponding subsets.
13. Do steps 6 to 8 with GMF 'ap'-index data.
14. Create a sliding window function which computes Rosenthal's effect size for the MMI, practice and control run data and the average values for the corresponding GMF 'ap'-index data.
15. Adjust the window width of the sliding window function (25°, 30°, 35°, 40°, 45°, and 50°).
16. Move the sliding window in steps of 0.5° over the entire lunar phase synchronously over all subsets.
17. Correlate the effect size values with the average GMF 'ap'-index data for every sliding window step.
18. Plot the Pearson correlation values.

Tests with split data

A split data test checks that the effect found is not an artifact of the precise sequence of recorded data. Therefore, the relevant data set was divided randomly into two samples. If there is a non-random effect in the data it will be present in both data sets. Two empty data samples were generated for the data split. A pseudo random function with current time as a random seed returned a number between 0 and 1 for every data split. If the pseudo random number was > 0.5, data were saved in sample #1, otherwise in sample #2. In this way, we obtained two samples which are independent of each other. The window width of the sliding window function was set to 36.6°, which is similar to the range of the full moon interval defined by Radin. Both samples were correlated with the GMF 'ap'-index data set. For both data sets we ran the sliding window function across each data set in steps of 0.5°, resulting in two rows of 16 average GMF 'ap'-index data values and 16 corresponding effect size per trial values for every 0.5° step. For this pair of 16 values each, Pearson's r with N = 16 was computed, so we got 720 Pearson correlation values for one sliding window function test across the whole lunar cycle.

Results

The overall results for the total of 199,632 'for the record' experiments yielded a per trial effect size of .49997 (where 0.5 is expected by chance in this test) which is equivalent to a nonsignificant *z* score of -0.82. The control runs yielded a per trial effect size of .50038 which is equivalent to a nonsignificant *z* score of 0.82, and the practice runs yielded a per trial effect size of .50001 which is equivalent to a nonsignificant *z* score of 0.32. But we have already seen that Radin found a full moon effect in the casino data with a higher payout rate within the days of the full moon, and we can assume that the total payout rate of the casino might not be significantly above the mean chance expectation. Otherwise, casino holders might worry about their future. If Radin's full moon effect is based on MMI, we can expect a similar effect in the Fourmilab Retropsychokinesis data.

Control run and practice data correlation results

A test with randomly split data was executed for the 36.6° full moon interval (161.7° to 198.3°). The control and practice run data yielded non-significant correlation values. For the first control run sample *r*(14) was -.15, and for the second sample *r*(14) was -.09. For the first practice data sample *r*(14) was .01, and for the second sample *r*(14) was .13. Thus, we might assume that there is no periodical bias in the data which could generate the full moon effect found in the MMI data. Figure 3 shows the results for the control run data.

The control runs showed a random walk of the curves. No full moon effect and no other coherent effects are visible in this figure. This seems to indicate that the HotBits hardware random number generator is working well for our purposes here.

MMI 'for the record' data correlation results

The test with randomly split data yielded the highest correlation value for MMI data. The first sample data yielded *r*(14) = -.62 with *p* = .005, one-tailed, and the second sample data yielded *r*(14) = -.54 with *p* = .015, one-tailed. Both datasets showed strong negative correlations which might indicate that there is a non-random trend in the data. Figure 4 shows the results for the MMI 'for the record' data. By visual inspection, both curves showed a similar peak in the full moon interval which reached significance. This might indicate that there is a non-random trend in both datasets which is lunar phase dependent.

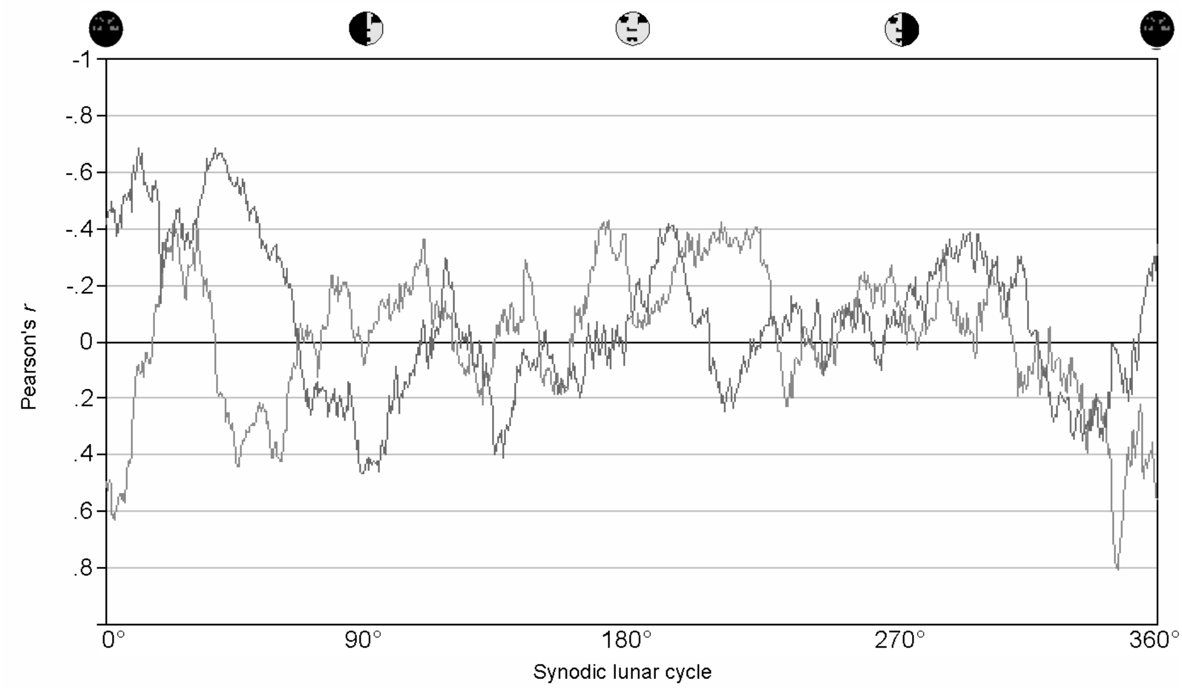


Figure 3. Smoothed epoch analysis for the split control run data vs. GMF 'ap'-index data. Data set A: grey line; data set B: black line.

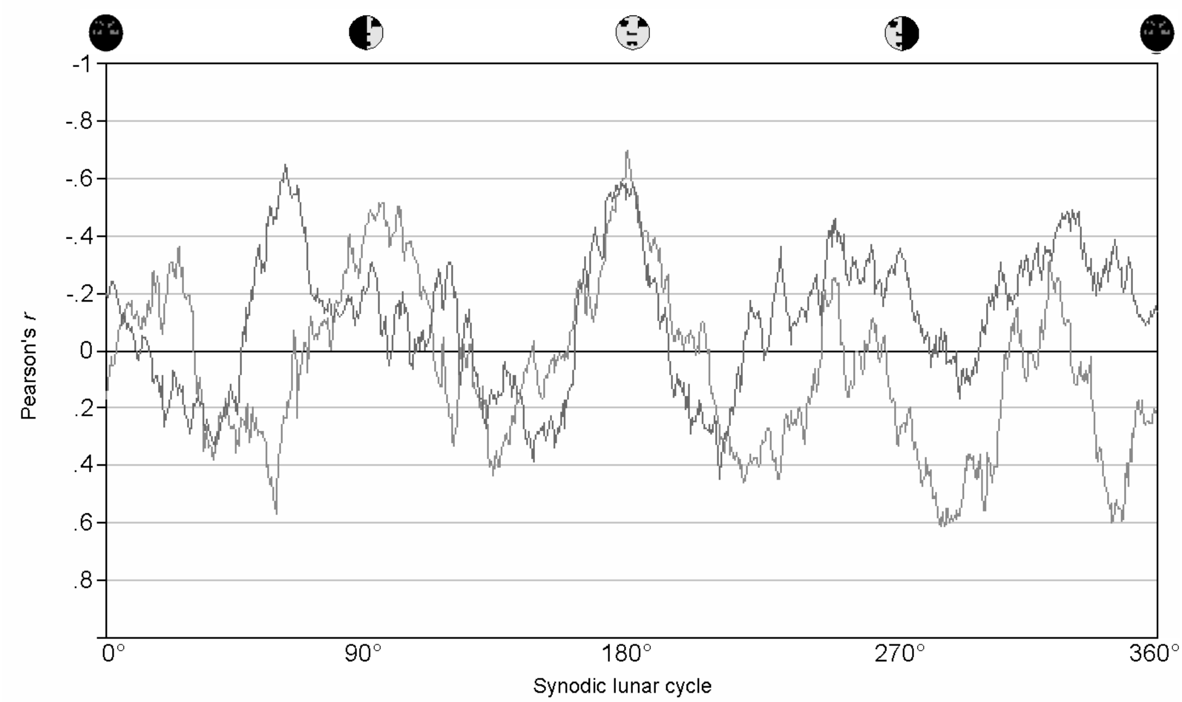


Figure 4. Smoothed epoch analysis for the split MMI 'for the record' experiment data vs. GMF 'ap'-index data. Data set A: grey line; data set B: black line.

Correlation results of non-split data & with different scanning window widths

We wished to plot the shape of the full moon effect and therefore the entire non-split MMI database was scanned with different sliding window widths of 25°, 30°, 35°, 40°, 45°, and 50°, starting at 0° through 360°, moving in steps of 0.5° for each correlation value. Figure 5 shows the results:

We found a sharply-defined full moon correlation effect which is visible for all scanning window widths. The peak effect was found at 3.5° before full moon with *r*(14) = .804 and 40° window width. The shape of the full moon peak resembled an imprint of Earth's magnetotail in the correlation data. A comparison with other Moon phase intervals showed that this strong negative correlation only exists for the full moon interval.

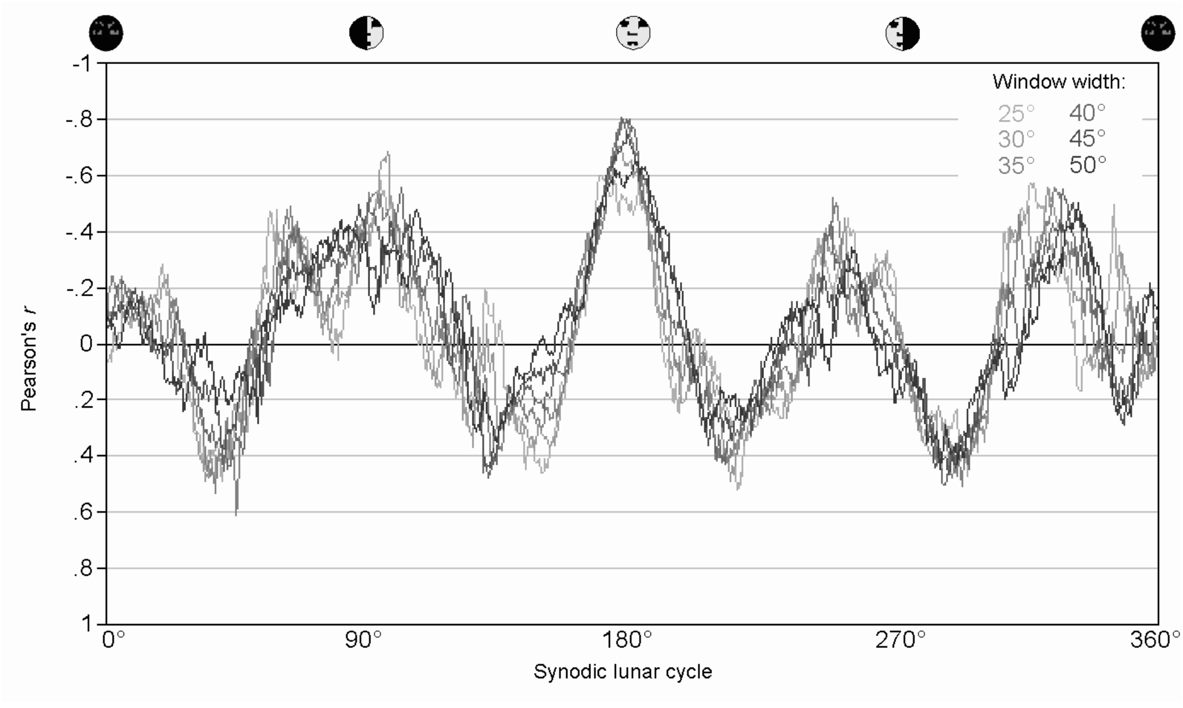


Figure 5. Epoch analysis for the retro-PK 'for the record' experiment data vs. GMF 'ap'-index data with different scanning window widths.

Table 1 lists the correlation values of the full moon MMI data for different window widths, centered at full moon and for comparison, the correlation values for the remaining non-full moon data.

Table 1. Pearson Correlation values for 180° lunar phase (full moon) and the remaining non-full moon data with $n = 16$.

Window width	**Pearson correlation of full moon data**	**Pearson correlation of non-full moon data**
25°	-.46*	-.030
30°	-.59**	-.026
35°	-.69**	-.008
40°	-.76***	.008
45°	-.69**	.008
50°	-.62**	.066

* $p < .05$; ** $p < .01$; *** $p < .001$ (all probabilities one-tailed)

All MMI data of the full moon intervals yielded significant results. For Radin's exact full moon interval of 161.7° to 198.3° lunar phase, we got $r(14) = -.74$ with $p = .0005$, one-tailed due to Radin's prediction of a negative correlation. The subset with the lowest average GMF 'ap'-index data value of 8.41 yielded a MMI per trial effect size of .5008 with a standard error of .00043, and the subset with the highest average GMF 'ap'-index data value of 29.1 yielded a MMI per trial effect size of .4993 with a standard error of .0004.

With a window width of 25° we found the initial point of the full moon correlation effect at 153° lunar phase and the end point at 198° lunar phase. With a window width of 50°, the initial point is at the position of 155° lunar phase and the end of the full moon effect at 208° lunar phase. This is the range in which we expect the Moon's transit through the magnetotail of the Earth. The Moon might pass the tail magnetopause between 198° and 208° (mean = 203°) and between 153° and 155° (mean = 154°) lunar phase, based on the observations here. The resulting average width of the Earth's magnetotail in the lunar orbit might be 49° lunar phase or a time interval of 4 days.

The center of the magnetotail in the lunar orbit might not be the exact full moon time, but 6 hours and 53 minutes before the exact full moon time. This might be in accordance with the expected shape of Earth's magnetotail due to the fact that the magnetotail is blown over by the interplanetary magnetic field.

Discussion

The tests in the current study seem to confirm previous observations of MMI-lunar relationship dependent on geomagnetic activity. If we assume that medical treatment such as distant healing or anomalous healing depend on MMI, then we may expect that these medical treatments might also be modulated by lunar-GMF effects, especially at full moon times of low GMF activity. But this is contrary to other findings. For example, Palmer, Baumann and Simmonds (2006) studied the magnetic influence on MMI effects whilst influencing the hemolysis of red blood cells, and wrote with respect to their study: "The most sobering implication of our data ... derives from the evidence of hemolysis acceleration. Translated into healing terms, this means that healers could unintentionally 'mis-direct' their PK to make an illness worse rather than better. ... If the finding with the Ap index continues to hold up, it may suggest that healing should be performed on a day following that when the global GMF is relatively high." The findings in the current study suggest that this statement is too general. The actual situation seems to be more complex. Healers might perform better on days of low GMF activity in full moon times with deep plasmasheet encounters. But if they try to heal at full moon times of extreme high GMF activity, they could make an illness worse rather than better because MMI might be active, but working against their intention.

An important observation in this analysis might be the difference between MMI 'for the record' data and the practice data for the full moon interval. While a remarkable full moon effect was found in the MMI 'for the record' data, it is completely absent in the practice run data. Due to the fact that practice and MMI 'for the record' data were taken from the same random source, one might assume that the difference might be due merely to psychological factors. The missing full moon effect in the practice data might indicate that a non-serious attempt to "influence" the output of the RNG might fail to produce a full moon effect. It might be dependent on the strength or gravity of intent. Practice alone does not yield a result. Translated into healing terms, this means that praying, wishing and healing without a serious intent might not yield a result.

Another point is the accurate shape of the full-Moon effect. We don't know what happens when the Moon is passing through the magnetotail but we may assume that a physical parameter is influenced which has a strong effect on the human mind to influence matter without the use of any currently known type of physical mechanisms. The current findings suggest that MMI might not be an elusive and weak effect but rather a stronger and more reliable effect if we knew the crucial parameters and could apply them systematically in MMI research.

We also do not understand the exact structure of Earth's farther magnetosphere. What we have are reconstructions and theoretical models of the geomagnetic field as observed by spacecraft data (Tsyganenko, 1995). These spacecraft are located at a distance of 3 to 40 Earth radii maximum (Tsyganenko and Sitnov, 2007). If the sources of this MMI effect observed here are based upon interactions of the Moon with the Earth's magnetotail, then we have a well-formed imprint of the farther magnetosphere which, beyond the use of expensive spacecraft, give us empirical clarification about the location of the magnetopause and the shape of the farther magnetotail at a distance of about 60 Earth radii. It seems to be unique in parapsychological research that an unusual MMI effect might be able to operate as a low cost probe for geomagnetic tail research.

While the casino data in the Radin study were generated at the same time and in the same location where they had been "treated", the MMI data had already been generated at an unknown time before being used in the Fourmilab RetroPsychokinesis Project. They were delayed for hours, days, or perhaps weeks, and they were generated far away from the participants' location. In both cases, we noticed an unusual effect, phase locked with the synodic lunar phase. If the retroactive psychokinesis effect really exists (Schmidt, 1976) – and the lack of any reverse effect in control and practice data support this assumption – we would not expect an effect in the parts of the lunar phase where the random data had been generated beforehand, but in this part of the lunar phase where the test persons observe the data in the MMI experiments. This was exactly what we had found. This observation might provide evidence of an ability of the human consciousness to influence targets distributed in space and time without being limited by the constraints of the local bodily presence (Schmidt, 1981).

References

Bell B., and Defouw R.J. (1964). Concerning a lunar modulation of geomagnetic activity. *Journal of Geophysical Research, 69*, 3169-3174.

Bigg E.K. (1963). The influence of the Moon on Geomagnetic disturbances. *Journal of Geophysical Research*, *68*, 1409-1413.

Bösch H., Steinkamp F., and Boller E. (2006). Examining psychokinesis: The interaction of human intention with random number generators: A meta-analysis. *Psychological Bulletin*, *132*, 497-523.

Etzold E. (2000). Lunarperiodische und solarperiodische Einflüsse in Psychokineseversuchen [Lunar-periodic and solar-periodic effects in psychokinesis experiments]. *Grenzgebiete der Wissenschaft*, *49*, 149-174.

Etzold E. (2002). Sind die Daten der Fourmilab-Experimente mit der Mondphase korreliert? Ein Replikationsversuch [Are the data of the Fourmilab experiments correlated with Moon phase? An attempted replication]. *Zeitschrift für Anomalistik*, *2*, 76-90.

Etzold E. (2005). Solar-periodic Full Moon Effect in the Fourmilab RetroPsychokinesis Project Experiment Data: An Exploratory Study. *Journal of Parapsychology*, *69*, 233-261.

Gearhart, L., and Persinger M.A. (1986). Geophysical variables and behaviour: XXXIII. Onsets of historical and contemporary poltergeist episodes occurred with sudden increases in geomagnetic activity. *Perceptual and Motor Skills*, *62*, 463-466.

Ghione S., Mezzasalma L., Del Seppia C., and Papi F. (1998). Do geomagnetic disturbances of solar origin affect arterial blood pressure? *Journal of Human Hypertension*, *12*, 749–754.

Gobba F., Bravo G., Scaringi M., and Roccatto L. (2006). No association between occupational exposure to ELF magnetic field and urinary 6-sulfatoximelatonin in workers. *Bioelectromagnetics*, *27*, 667-73.

Hapgood M. (2007). Modelling long-term trends in lunar exposure to the Earth's plasmasheet. *Annales Geophysicae*, *25*, 2037-2044.

Heath P.R. (2003). *The PK Zone. A Cross-Cultural Review of Psychokinesis* (PK). Lincoln, iUniverse.

Jahn R., Dunne B., Dobyns Y., Nelson R., Bradish, G., Lettieri A., et al. (2000). Mind/machine interaction consortium: PortREG replication experiments. *Journal of Scientific Exploration*, *14*, 499-555.

Knott K. (1975). A lunar signature in the geomagnetic ap-index. *Earth, Moon, and Planets*, *14*, 49-57.

May E.C., Utts J.M., and Spottiswoode J.P. (1995). Decision Augmentation Theory: Toward a Model of Anomalous Mental Phenomena. *Journal of Parapsychology*, *59*, 195-220.

O'Connor R.P., and Persinger M.A. (1999). Geophysical variables and behavior: LXXXV. Sudden infant death, bands of geomagnetic activity, and pc1 (0.2 to 5 HZ) geomagnetic micropulsations. *Perceptual and Motor Skills*, *88*, 391-397.

Palmer J., Simmonds C.A., and Baumann S. (2006). Geomagnetic fields and the relationship between human intentionality and the hemolysis of red blood cells. *Journal of Parapsychology*, *70*, 275-301.

Persinger M.A. (1987). Geopsychology and geopsychopathology: mental processes and disorders associated with geochemical and geophysical factors. *Experientia*, *43*, 92-104.

Radin D.I. (1997). *The Conscious Universe. The scientific truth of psychic Phenomena*. San Francisco, HarperEdge.

Radin D.I., and Rebman J.M. (1998). Seeking Psi in the Casino. *Journal of the American Society for Psychical Research*, *62*, 193-219.

Radin D.I., Nelson R., Dobyns Y., and Houtkooper J. (2006). Reexamining Psychokinesis: Comment on Bösch, Steinkamp, and Boller. *Psychological Bulletin*, *132*, 529-532.

Schmidt H. (1976). PK effect on pre-recorded targets. *Journal of the American Society for Psychical Research*, *70*, 267-291.

Schmidt H. (1981). Can an effect precede its cause? *Foundations of Physics*, *8*, 463-480.

Schubert G., Sonett C.P., Smith B.F., Colburn D.S., and Schwartz K. (1975). Using the Moon to probe the geomagnetic tail lobe plasma. *Geophysical Research Letters*, *2*, 277-280.

Spottiswoode S.J.P. (1990). Geomagnetic activity and anomalous cognition: A preliminary report of new evidence. *Subtle Energies*, *1*, 91-102.

Sturrock P.A. (2004). Time-Series Analysis of a Catalog of UFO Events: Evidence of a Local-Sidereal-Time Modulation. *Journal of Scientific Exploration*, *18*, 399–419.

Sturrock P., and Spottiswoode S.J.P. (2007). Time-series power spectrum analysis of performance in free response anomalous cognition experiments. *Journal of Scientific Exploration*, *21*, 47-66.

Tsyganenko N.A. (1995). Modeling the Earth's magnetospheric magnetic field confined within a realistic magnetopause. *Journal of Geophysical Research*, *100*, 5599-5612.

Tsyganenko N.A., and Sitnov M.I. (2007). Magnetospheric configurations from a high-resolution data-based magnetic field model. *Journal of Geophysical Research*, *112*, A06225, doi:10.1029/2007JA012260.

Wilson B.W., Wright C.W., Morris J.E., Buschbom R.L., Brown D.P., Miller D.L., et al. (1990). Evidence of an effect of ELF electromagnetic fields on human pineal gland function. *Journal of Pineal Research*, *9*, 259-269.

European Journal of Parapsychology
Volume 23.2, pages 173-188

ISSN: 0168-7263

News Events, False Memory and Paranormal Belief

Neil Dagnall, Andrew Parker and Gary Munley
Department of Psychology & Social Change, Manchester Metropolitan University

Abstract

The current study explores Wilson and French's (2006) finding that believers in the paranormal are more susceptible to false memories than non-believers. Seventy participants completed a News Coverage Questionnaire concerning personal memories about first seeing dramatic news events. The events included four news items that are known to have been captured on film and two items for which there is no known footage (the Princess Diana car crash and the Bali Bombing). Presentation order was found to affect the reporting of false memories. Endorsement of the Diana false memory item positively correlated with scores on the Australian Sheep-Goat Scale and the Dissocative Experience Scale when the item appeared first but not when it appeared fifth. Believers in the paranormal were more susceptible to false memory than non-believers when the critical item was answered without first recalling actual television footage. However, recalling previous news footage facilitated endorsement of the false memory item independent of paranormal belief and level of dissociativity.

Introduction

False memory has frequently been investigated using the "crashing paradigm" (Crombag, Wagenaar, & van Koppen, 1996). This involves examining participants' willingness to report memories of a highly emotional and consequential event that they could not have witnessed (Crombag et al., 1996; Granhag, Strömwall, & Billings, 2003; Ost, Vrij, Costall, & Bull, 2002; Wilson & French, 2006). This approach

Correspondence details: Neil Dagnall, Department of Psychology & Social Change, Manchester Metropolitan University, Elizabeth Gaskell Campus, Hathersage Road, Manchester, M13 0JA
Email: N.Dagnall@mmu.ac.uk

was pioneered by Crombag, et al. (1996) and is based on the notion that participants will be influenced by a critical question, which falsely suggests that actual film footage exists (Crombag et al., 1996; Loftus, 1975).

Crombag et al. (1996) questioned participants 10 months after the crash of an E1-A1 Boeing 747 into a block of flats in Amsterdam (October, 1992). Although, there was no actual footage of the plane crash the event received extensive media coverage. When asked, "Did you see the television film of the moment the plane hit the apartment building?", 55% of the participants reported that they had seen non-existent footage. In addition to this, 82% of participants provided details about a non-existent fire. In a second study Crombag et al. (1996) replicated these results. When participants were asked to provide highly specific details about the crash (e.g., the angle at which the plane crashed into the building) 66% of participants claimed to have seen the footage and the majority were able to provide detailed responses to the questions.

Other studies testing recollection of high profile events, for which there is no known footage, have produced similar results: the Diana, Princess of Wales, car crash (Ost, Vrij, Costall, & Bull, 2002); the sinking of the Estonia ferry, in which almost 900 lives were lost (Granhag et al., 2003); and the assassination of Dutch politician Pim Fortuyn (Jelicic, Smeets, Peters, Candel, Horselenberg, & Merckelbach, 2006). Each of these studies produced impressive results. Ost et al. (2002) using the Diana Princess of Wales car crash found that 44% of participants were willing to report that they had seen actual television footage of the crash. Granhag et al. (2003) noted that 38% of participants in study 1, and 55% in study 2, reported seeing footage of the sinking of the Estonia ferry. Jelicic et al. (2006) found that 63% of participants claimed to have seen non-existent footage of the assassination of Pim Fortuyn, although only 23% could provide details. It is evident from these studies that participants will frequently report having witnessed events they could not possibly have seen; events for which no known video footage exists.

Wilson and French (2006) using a modified version of the crashing memories paradigm examined the relationship between false memory and paranormal belief/experience. They contended that research in this area was important because variables associated with susceptibility to false memories (e.g., dissociativity, fantasy proneness, absorption) had also been found to correlate with paranormal belief/experience (Wilson and French, 2006).

To test this, Wilson and French (2006) asked participants to complete a News Coverage Questionnaire (NCQ), the Australian Sheep–Goat Scale (ASGS) (Thalbourne, 1995), Dissociative Experiences Scale (DES) (Bernstein & Putnam, 1986), and the Anomalous Experiences Inventory (Kumar, Pekala, & Gallagher, 1994). The NCQ contained four events that are known to have been captured on film (the attack on the World Trade Centre, the Hillsborough football stadium disaster, the toppling of the Saddam Hussein statue in Basra, and the Challenger space shuttle disaster) and one item concerning non-existent footage (the bombing of a nightclub in Bali). Participants were asked specific questions about each of the events (where they were, who they were with, and where they saw the footage). Questions were also asked about the footage (e.g., was it in colour or black and white, was there a commentary, and what was the picture quality like). Wilson and French (2006) found that 36% of respondents falsely claimed to have seen footage of the Bali bombing. Participants reporting false memories were found to score higher on the ASGS, the DES, and on the Belief, Experience and Ability subscales of the Anomalous Experiences Inventory than participants who did not report seeing the non-existent footage. On the basis of their findings Wilson and French (2006) concluded that believers in the paranormal may be more susceptible to false memories than non-believers. This assertion, whilst logical, has yet to be firmly established and will be examined within the current paper.

The present study

Whilst Wilson and French (2006) found support for the hypothesis that believers in the paranormal may be more susceptible to false memories than non-believers, other studies exploring the relationship between susceptibility to false memories and the tendency to report anomalous experiences have achieved only limited success (French, 2003; French & Wilson, 2006). French, in his review of eyewitness testimony and false memories for reports of anomalous experiences (French, 2003; Wilson & French, 2006), suggests that this may be because studies have used inappropriate techniques to measure susceptibility to the “type” of false memories expected to correlate with the tendency to report anomalous experiences.

Particularly, French suggests that techniques attempting to produce detailed false memories for entire episodes (e.g., Loftus & Pickrell, 1995) may be more successful than techniques which attempt to produce false memories for individual words (e.g., Roediger & McDermott, 1995; word lists). This notion was supported by Wilson and French's (2006) findings with news events.

The problem with the Wilson and French (2006) study is that endorsement of the false memory item may be conflated by the prior exposure of participants to news events for which footage exists. In the Wilson and French (2006) paper the critical Bali bombing event was presented after two events for which there was actual footage. Presenting the critical scenario in this way may encourage participants to make reporting errors. Remembering actual footage may bias participants' responses when faced with an event for which no footage exists; participants may assume the event was shown on television, or that their memory of the event was based on television coverage. If participants are being influenced in this way, it would be predicted that item order would affect the reporting of false memories. Under conditions where participants are first asked to recall real footage it seems plausible that they will be more likely to misreport recalling non-existent news footage. Indeed, Wilson and French (2006) hypothesized that this may be an important factor requiring further research. Noting that the level of false memory endorsement was higher in their study than comparative studies (e.g., Jelicic et al., 2006), they speculated that this was because previous studies used a single event; Wilson and French (2006) presented the false memory amongst four other events that had been filmed. For this reason the current study assesses whether recalling footage of actual events influences participants' false memory responses.

This is achieved in the present paper by manipulating news event order. The critical event will be presented first and last (fifth) within an adapted version of the NCQ. The story selected for this purpose is the fatal Paris car crash involving Diana, the Princess of Wales, and her companion Dodi Fayed. This item has been used in other similar studies, where participants have been found to wrongly report seeing actual crash footage (Ost et al., 2002). This scenario was selected over the Bali bombing because it received extensive coverage and the story has remained in the news since the episode occurred. The Bali bombing, whilst an important happening, has received less prominent coverage since the event, consequently, it was felt that participants would be less familiar with the Bali bombing. In order to allow direct comparison with Wilson and French (2006), the Bali bombing will still be used in the current study; it is presented at the end of the experiment to determine if levels of false memory are similar to those reported by Wilson and French.

Wilson and French (2006) hypothesized that item order would influence the reporting of false memory. In particular they claimed that recalling real news footage in preceding items would prime or activate a generic schema for news coverage. Once activated, this news schema influences the encoding and comprehension of non-existent news items and leads to that information appearing more real. Under such conditions, the individual is more likely to erroneously assert that such footage actually existed. Thus it is predicted that presenting the Diana false memory item last rather than first will increase false memory item endorsement and that presenting the Bali bombing question at the end of the current study will result in a higher level of endorsement than that reported in the Wilson and French (2006) study. In addition to this the present study will assess whether item order affects the relationship between paranormal belief, dissociative experience and false memory, In line with Wilson and French (2006) it is predicted that level of paranormal belief and dissociative experience will be associated with false memory endorsement and the rating of false memory characteristics. The major contribution of the present study will be to assess whether the relationship between paranormal belief, dissociative experience and false memory is qualified by item order.

Method

Participants

Seventy respondents participated in this study. There were 25 males and 45 females. The mean age was 30.63 years (SD = 12.51) with a range of 16–67 years. Participants included undergraduates and employees from the Manchester Metropolitan University and volunteers from the wider community.

Materials

Participants were asked to complete in the following order: an adapted version of the News Coverage Questionnaire (NCQ) (Wilson

& French, 2006); the Australian Sheep–Goat Scale (ASG) (Thalbourne, 1995); and the Dissociative Experiences Scale (DES) (Bernstein & Putnam, 1986). Questionnaire order was the same for all participants.

The News Coverage Questionnaire was designed by Wilson and French (2006). The questionnaire contains five news items; four events that are known to have been captured on film (the 2001 attack on the World Trade Centre; the 1989 Hillsborough football stadium disaster; the 2003 toppling of the Saddam Hussein statue in Basra; and the 1986 Challenger space shuttle disaster) and one item concerning non-existent footage (the 2002 Bali nightclub bombing). Participants were asked to indicate: whether they recall first seeing the event; where they saw the event; and who they were with at the time. Questions were asked about the actual news item footage: whether the clip was in colour or black and white; whether there was an audible commentary on the footage; and how participants would rate the picture quality. These questions were intended to ensure that participants reporting false memories based their report upon a specific memory of viewing, rather than just recalling the circumstances under which they first heard the news being reported.

In the current study an additional item was added. This concerned watching non-existent footage of the 1997 car crash involving Princess Diana. This item was selected because participants have been found to often report seeing the actual video footage of the crash in which Diana, Princess of Wales, and Dodi Fayed died, even though no footage exists (Ost et al., 2002). This news event shares many important features with the El Al air crash used in Crombag et al.'s (1996) study: the media extensively covered the story; television footage appeared soon after the event; the crash was the main news story; the story was told again and again illustrating what footage was actually available; people knew many details about the event without actually seeing it; and a simulation of the event was shown on some news stations.

In order to allow the current study to assess the effect of item order on false memory, the order of the items for which there is no known television footage was manipulated. The Diana news event appeared in either first or fifth; that was before or after the events for which there was known footage (World Trade Centre, Hillsborough, Space Shuttle Challenger and Basra). The second critical item, the Bali Bombing was presented at the end of the study to allow direct comparison with Wilson and French (2006). The NCQ was followed by the Australian Sheep–Goat Scale (ASG) and DES.

The Australian Sheep–Goat Scale (ASGS) measures belief in, and alleged experience of, the paranormal. It consists of 18 items that relate to three core concepts of parapsychology: life after death, psychokinesis, and extrasensory perception. The response options are: False (scored as zero), "?" (Don't know: scored as 1 point), and True (scored as 2 points). The scale has a range from zero to 36, higher scores indicating higher levels of belief and experience. The ASGS has proven reliability and validity (Thalbourne, 1995; Thalbourne and Delin, 1993).

The Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986) is a 28-item self-report measure that assesses the frequency of dissociative experiences on a response scale that starts at 0% and increases by 10% increments up to 100% (Carlson & Putnam, 1993). Responses across all items are averaged to obtain a mean DES score (range 0–100), with higher scores indicating higher levels of dissociative symptoms.

Procedure

Participants were informed that the current study was investigating memory for news events. Following completion of the questionnaire participants were debriefed. They were informed that there was no actual footage of the Diana crash and the Bali bombing.

Results

Both the Australian Sheep–Goat Scale (ASGS) (α = .92) and Dissociative Experiences Scale (DES) (α =.90) were found to have excellent internal reliability. Of the 70 participants who completed the questionnaire, 68 (97%) reported that they had seen the video footage of the attack on the World Trade Centre, 26 (37%) reported having seen the video footage of the Hillsborough football disaster in 1989, 65 (93%) recalled video footage of the statue of Saddam Hussein being torn down in Basra in 2003, and 22 (31%) recalled the video footage of the Challenger space shuttle disaster in 1986. Regarding the non-existent video footage, 57 (81%) participants claimed to have seen the car crash involving Princess Diana; endorsement was higher on this item than on two of the events for which actual video footage exists

(Hillsborough and the space shuttle Challenger). Table 1 contains the percentage claiming recall for the present study and for Wilson and French (2006).

Table 1: Percentage Claiming to Recall Television Footage of Specific News Events in the Current Study and in the Wilson and French (2006) Study

	Percentage claiming recall	
News story	Current study	Wilson & French (2006)
World Trade Center	97	97
Hillsborough	37	30
Basra	93	72
Challenger	31	40
Bali	53	36
Diana	81	*NA*

To determine whether endorsement of the Diana crash (false memory item) was influenced by scenario presentation order (presented first vs. fifth) chi-square was used. A significant association was found between presentation order and item endorsement, χ^2 = 5.14, df = 1, p = .023, Cohen's d = 0.56. In particular, 77% of correct rejections occurred when the Diana event was presented first this fell to 23% when the event was presented fifth (last), χ^2 = 3.77, df = 1, p = .052, Cohen's d = 0.54. Endorsement of the Diana crash was further studied by examining false alarms, the number of incorrect endorsements produced, This found that presentation order had no significant effect on the number of false alarms produced, χ^2 = 1.42, df =1. p > .05, Cohen's d = 0.32; the trend towards fewer false alarms 42% when presented first, compared to 58% when presented fifth was, however, in the predicted direction. Finally, the difference between correct rejections and false alarms as a function of presentation order was analysed. True memory was found to be higher when the Diana crash was presented first than when it was presented fifth, χ^2 = 5.82, df =1. p = .016, Cohen's d = 0.78.

Of the 57 participants who claimed to have seen non-existent footage of the Diana crash, 88% were willing to state where they were when they saw the footage, 71% recalled who they were with at the time, and 51% recalled what TV channel/website they saw the footage on (95% claiming it was in colour). The picture quality of the footage, on a 7-point scale was 5.12 (SD = 1.47) and the mean clarity of memory was 5.00 (SD = 1.82).

The mean clarity rating for events with actual footage was compared with the mean rating for the Diana non-existent television footage. Participants who endorsed having seen the Diana footage (M = 5.00) claimed to recall it as clearly as they did actual footage (M = 5.02); t (56) = - 0.09, $p > .05$.

A median split was performed on ASGS scores to produce high (above median) and low (below median) conditions. A 2 (Level of Paranormal Belief: high vs. low) x 2 (Presentation Order: first, fifth) factor between subjects ANOVA was then conducted to determine whether variable level (high vs. low) and presentation order affected participants' clarity ratings of the non-existent Diana car crash footage (see Table 2).

Table 2: Clarity Ratings of the Non-Existent Diana Car Crash Footage as a Function of level of Paranormal Belief

	Level of paranormal belief					
	Below Median		Above Median		Overall	
Presentation order	*M*	*SD*	*M*	*SD*	*M*	*SD*
First	4.80	1.62	4.93	1.33	4.88	1.42
Fifth	3.86	2.18	6.00	1.49	5.09	2.08
Overall	4.25	1.94	5.55	1.50		

A significant main effect was noted for Level of Paranormal Belief, F(1, 53) = 6.24, $p = .016$, $\eta p^2 = 0.11$. Participants high in paranormal belief (M = 5.55, SD = 1.50) rated the clarity of the non-existent footage higher than participants low in paranormal belief (M = 4.25, SD = 1.99).

No significant effect was found for Presentation Order, F(1, 53) = .20, $p > .05$. The interaction between Level of Paranormal Belief and Presentation Order, F(1, 53) = 4.91, $p = .031$, $\eta p^2 = 0.09$ was significant.

Simple main effect analysis on each level of Presentation Order showed that Level of Paranormal Belief had no effect on clarity rating when the Diana car crash question was presented first, t(22) = 0.21, $p > .05$, however, it did affect ratings when the Diana car crash question was presented fifth. Participants high in paranormal belief (M = 6.00, SD = 1.48) were found to rate clarity for the non-existent footage higher than those low in paranormal belief (M = 3.86, SD = 2.18; t(31) = 3.36, $p = .002$, Cohen's $d = 1.21$).

A median split was performed on DES scores to produce high (above median) and low (below median) conditions. A 2 (Level of

Dissociativity: high vs. low) x 2 (Presentation Order: first, fifth) ANOVA was then conducted to determine whether DES score (high vs. low) and presentation order affected participants' clarity ratings of the non-existent Diana crash footage (see Table 3).

Table 3: Clarity Ratings of the Non-Existent Diana Car Crash Footage as a Function of level of Dissociativity

	Level of dissociativity					
	Below Median		Above Median		Overall	
Presentation order	*M*	*SD*	*M*	*SD*	*M*	*SD*
First	5.18	1.66	4.61	1.19	4.88	1.42
Fifth	4.31	2.32	5.60	1.79	5.09	2.08
Overall	4.71	2.05	5.21	1.64		

No significant main effect was found for Level of Dissociativity $F(1, 53) = 0.56$, $p > .05$. No significant main effect was noted for Presentation Order, $F(1, 53) = 0.13$, $p > .05$. The interaction between Level of Dissociativity and Presentation Order $F(1, 53) = 3.65$, $p > .05$ was not significant.

A series of correlations were conducted to investigate the relationship between Australian Sheep–Goat Scale (ASGS) and Dissociative Experiences Scale (DES) scores, and whether or not participants claimed to have seen television footage of the car crash involving Princess Diana. ASGS was found to positively correlate with participants' endorsement of the non-existent Diana television footage ($r = .38$, $n = 70$, $p = .001$). A significant correlation was found between ASGS and the clarity rating of the non existent Diana memory ($r = .35$, $n = 57$, $p = .007$). DES was found to positively correlate with participants' endorsement of the non-existent Diana television footage ($r = .28$, $n = 70$, $p = .017$). No significant relationship was found between DES and the clarity rating of the non existent Diana memory ($r = .17$, $n = 57$, $p > .05$). The DES and ASGS were also found to correlate ($r = .37$, $n = 70$, $p = .002$).

Further correlations were conducted on participants' responses when the Diana item was presented first and fifth. When the Diana item appeared first positive correlations were found between false memory endorsement and paranormal belief ($r = .43$, $n = 34$, $p = .011$) and between false memory endorsement and DES ($r = .39$, $n = 34$, $p = .022$). When the Diana item appeared fifth no significant relationship was found between false memory endorsement and paranormal belief (r = .28, n = 36, $p > .05$) and between false memory endorsement and DES (r = .16, n = 36, $p > .05$).

These results indicate a significant positive correlation between false memory and belief in the paranormal and DES when the false memory item is not preceded by events for which footage exists. However, no relationship was found between false memory and belief in the paranormal and DES when the false memory item was presented after events for which there was known footage.

Endorsement of the non-existent Bali bombing footage was presented at the end of the questionnaire in the current study to investigate whether presentation order affected the proportion of correct and incorrect responses. The data from this paper, where the event was presented last was compared with that of Wilson and French (2006); Wilson and French (2006) always presented the Bali item third of five. As predicted presenting the Bali item last produced a higher percentage of false endorsements (53% compared to 36% in the Wilson & French, 2006 study), $\chi^2 = 4.77$, df = 1, $p = .029$, Cohen's d = 0.54. Endorsement of the Bali false memory item may have been higher in the current study than the French and Wilson (2006) study because participants were influenced by previously recalling actual footage (World Trade Centre, Hillsborough, Basra, etc.)[1].

Discussion

Consistent with previous studies a substantial proportion of participants erroneously reported seeing non-existent news footage of critical events (crash of Diana and Bali bombing) (Crombag et al.,1996; Granhag et al., 2003; Ost et al., 2002; Smeets et al., 2006; Wilson & French, 2006). The current paper investigated the hypothesis that the reporting of false memories would be affected by presentation order as predicted by Wilson and French (2006). It was expected that a higher proportion of participants would report seeing the Diana crash when the critical question was presented fifth rather than first, and that presenting the Bali bombing question at the end of the current study would result in a higher level of endorsement than that reported in the Wilson and French (2006) study. Both of these predictions were supported: 77% of correct rejections occurred when the Diana item was presented first; and higher levels of endorsement were produced for

[1] Whilst it is unusual to combine data in this way, studies in this area frequently use percentage comparisons to support/refute points (e.g., Smeets, et al., 2006; Wilson & French, 2006).

the Bali item than were found in the Wilson & French (2006) study (53% vs. 36%).

The influence of serial order is explained by Wilson and French (2006) as due to the activation of a generic news schema. Presumably, once primed, the news schema provides a basis for the encoding and comprehension of incoming information. Non-existent footage is thus assimilated into this schema. When probed about the details of this footage, the generic schema is reactivated and false memories of non-existent events are the result. Interestingly, clarity ratings also increased when the critical item was presented fifth, at least for those high in paranormal belief. It would appear that these individuals had more vivid false recollections of the details of the non-existent event.

Exactly why this was found in terms of the news schema hypothesis is difficult to say. Schematic information, by its very nature, is abstracted away from specific perceptual or episodic details that may influence ratings of clarity (Alba & Hasher, 1983). Indeed, memory for schematic information is *less* likely to be associated with clear and vivid recollections (Lampinen, Faires, Neuschatz, & Toglia, 2000). If this is true, then the news schema hypothesis of Wilson and French (2006) may not provide a full, account of the effects observed here. Instead, perhaps one explanation is that individuals high in paranormal belief are more prone to incorporating perceptual or episodic features form real events into their false memories. In other words, they are more likely to "import" details from studied events into non-studied events making them seem more likely and real (e.g., Lyle & Johnson, 2006). Another possibility is that they adopt a more liberal criterion for assessing the degree of clarity and detail retrieved. Both these ideas warrant further investigation.

Previous research has also found that those score high on measures of paranormal belief also demonstrate reality monitoring deficits (Irwin, 2004). As false memories can also arise when individuals fail to accurately engage in reality monitoring, then false recollections of the Diana crash may also be due to reality monitoring failures in those endorsing high levels of paranormal belief.

The results of the current study indicate that participants are more likely to wrongly report seeing television coverage of an event when the critical false memory item is preceded by events for which actual footage exists. This supports the notion that remembering actual footage biases participants' responses; participants may assume the "critical event" was shown on TV, or that their memory of the event was based on television coverage (Wilson & French, 2006). Participants scoring above the median on the ASGS were found to rate the clarity of the non-existent Diana footage higher than those scoring below the median on the ASGS when the Diana question was presented fifth. No difference was evident when the Diana question was presented first.

In the current study overall endorsement of the non-existent Diana television footage was higher than that reported by Ost et al. (2002); 81% vs. 44%. This may be partially explained by the use of Wilson and French's (2006) multi-event paradigm, but may also be due to other factors: the continued news coverage since the event; recent refocusing (Stevens Inquiry, 2006); and the forthcoming tenth anniversary coverage (August 2007). This notion is consistent with the imagination inflation effect (Garry et al., 1996). Whereby, intense and sustained media coverage is likely to encourage people to imagine details of an event and motivate the belief that footage has appeared on television.

The overall clarity rating for false memory in the current study was moderately high similar to that produced by Wilson and French (2006) (5.00 compared to 4.25: on a 7-point scale). As in the Wilson and French (2006) study the lowest clarity ratings were produced for the Hillsborough and Challenger disasters. This is perhaps not surprising as these are the two oldest events. Interestingly, for participants who "wrongly" recalled seeing the Diana crash there was no difference between their false memory clarity rating and that for the real events (5.00 vs. 5.02). This finding is in line with previous research (e.g., Porter, Yuille, & Lehman, 1999) that has demonstrated that fabricated memories are sometimes rated as vivid and detailed as real memories (Smeets et al., 2006). In addition to this, a high proportion of participants provided details to accompany their endorsement of the Diana item.

In support of Wilson and French (2006), the current study found positive correlations between endorsement of the Diana item and scores on the Australian Sheep–Goat Scale (ASGS) and Dissociative Experiences Scale (DES). These findings are consistent with previous studies, which have reported relationships between false memory and psychological correlates of paranormal belief (e.g., fantasy proneness: Irwin, 1990; hypnotic suggestibility: Kumar & Pekala, 2001; absorption: Irwin, 1985; dissociativity: Wolfradt, 1997) (French & Wilson, 2006).

A major finding in the current study was that the relationship between false memory and scores on the AGS and DES was qualified by item order. A medium positive correlation was found between belief in the paranormal (ASGS), DES, and false memory for "spontaneous" endorsement of the Diana false memory item. However, this relationship was weakened when endorsement was induced by thinking about actual television footage of news events. The fact that endorsement of the false memory item was affected by the presentation of preceding events, for which there is known footage, suggests that recalling actual footage influenced responses on the critical Diana item independent of paranormal belief and dissociativity.

This finding is consistent with Smeets et al. (2006), who demonstrated that reporting of non-existent footage was shaped by misleading and/or ambiguous questions, and that the endorsement rate of footage depends on the way in which participants are questioned. Further to this, they identified the suggestibility of the participant as being important (Smeets et al., 2006). Similarly, studies have demonstrated that verbal suggestion can affect participants' perception of paranormal events, such as séances (Wiseman, Greening & Smith, 2003) and key bending (Wiseman & Greening, 2005). Like these previous studies the current study is limited by the fact that it fails to explain the relationship between false memory, paranormal and psychological variables; it is unclear whether the questions affect perception, memory, or whether participants are influenced by demand characteristics (French & Wilson, 2006).

The use of the crashing paradigm to assess false memory is not without criticism. The main problem being that it does not necessarily follow that participants claiming to have seen non-existent news footage have actually created false memories (Smeets, Merckelbach, Horselenberg, & Jelicic, 2005). Statement endorsement may be influenced by myriad cognitive and social factors (Smeets et al., 2006): participants may provide socially desirable answers; they may misinterpret the critical questions (refer to the aftermath rather than the actual event); use general knowledge heuristics; or make source monitoring errors rather than create false memories. Subsequent studies are required to assess the reasons for false memory endorsement.

In addition to this the current multi-event paradigm has the problem of being unable to ensure that memory for actual footage is accurate. This is not important to the degree that the current study is exploring whether participants endorse having seen non-existent footage. However, the fact that memories for events for which there is known footage may be untrue does bring into question the validity of subjective ratings. One obvious way of overcoming this problem is to show participants a series of news clips and then after delay, ask them about the events alongside events for which there is no known footage. This modification, whilst time consuming would ensure that participants had been exposed to the footage and that "true" memories had a clearly identifiable objective basis.

References

Alba J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93, 201–231.

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability and validity of a dissociation scale. *Journal of Nervous and Mental Diseases*, 74, 727–735.

Carlson, E. B., & Putnam, F. W. (1993). An update on the dissociative experiences scale. *Dissociation*, 6, 16-27.

Crombag, H. F. M., Wagenaar, W. A., & van Koppen, P. J. (1996). Crashing memories and the problem of 'source monitoring'. *Applied Cognitive Psychology*, 10, 95-104.

French, C. C. (2003). Fantastic memories: The relevance of research into eyewitness testimony and false memories for reports of anomalous experiences. *Journal of Consciousness Studies*, 10, 153-174.

French, C. C., & Wilson, K. (2006). Incredible memories: How accurate are reports of anomalous events? *European Journal of Parapsychology*, 21, 166-181.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: imagining a childhood event inflates confidence that it occurred. *Psychological Bulletin and Review*, 3, 208–214.

Granhag, P. A., Strömwall, L., & Billings, J. (2003). I'll never forget the sinking ferry: how social influence makes false memories surface. In Vanderhallen, M., Vervaeke, G., J. van Koppen, P., & Goethals, J. (eds.). *Much ado about crime: Chapters on psychology and law.* Belgium: Uitgeverij Politeia, 129-140.

Irwin, H., J. (1985). Parapsychological phenomena and the absorption domain. *Journal of the American Society for Psychical Research,* 79, 1-11.

Irwin, H., J. (1990). Fantasy proneness and paranormal beliefs. *Psychological Reports*, 66, 655-658.

Jelicic, M., Smeets, T., Peters, M. J. V., Candel, I., Horselenberg, R., & Merckelbach, H. (2006). Assassination of a controversial politician: remembering details from another non-existent film. *Applied Cognitive Psychology,* 20, 591-596.

Kumar, V .K., & Pekala, R. J. (2001). Relation of hypnosis-specific attitudes and behaviors to paranormal beliefs and experiences: a technical review. In Houran, J., & Lange, R. (eds.). *Haunting and poltergeists: multidisciplinary perspectives*. Jefferson, NC: McFarland & Co, 260–279.

Kumar, V. K., Pekala, R. J., & Gallagher, C. (1994). *The Anomalous Experiences Inventory*. Unpublished psychological test. West Chester, PA: West Chester University.

Lampinen, J. M., Faires, J. M., Neuschatz, J. S., & Toglia, M. P. (2000). Recollections of things schematic: the influence of scripts on recollective experience. *Applied Cognitive Psychology*, 14, 543-554.

Loftus, E. F. (1975). Leading questions and eyewitness reports. *Cognitive Psychology*, 7, 560-572.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, 25, 720-725.

Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory*, 14, 197-213.

Memon, A., & Young, M. (1997). Desperately seeking evidence: The recovered memory debate. *Legal and Criminological Psychology*, 2, 131-154.

Ost, J., Vrij A., Costall A., & Bull, R. (2002).Crashing memories and reality monitoring: distinguishing between perceptions, imaginings and false memories. *Applied Cognitive Psychology*, 16, 125–134.

Porter, S., Yuille, J.C., & Lehman, D. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior*, 23, 517-538.

Pekala, R. J., Kumar, V. K., & Marcano, G. (1995). Anomalous/paranormal experiences, hypnotic susceptibility, and dissociation. *Journal of the American Society for Psychical Research*, 89, 313–332.

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803-814.

Smeets, T., Jelicic, M., Peters, M. J. V., Candel, I., Horselenberg, R., & Merckelbach, H. (2006). 'Of course I remember seeing that film?' - How ambiguous questions generate crashing memories. *Applied Cognitive Psychology*, 20, 779-789.

Smeets, T., Merckelbach, H., Horselenberg, R., & Jelicic, M. (2005). Trying to recollect past events: confidence, beliefs, and memories. *Clinical Psychology Review*, 25, 917-934.

Thalbourne, M. A. (1995). Further studies of the measurement and correlates of belief in the paranormal. *Journal of the American Society for Psychical Research,* 89, 233–247.

Thalbourne M. A., & Delin, P.S., (1993). A new instrument for measuring the sheep–goat variable: its psychometric properties and factor structure, *Journal of the Society for Psychical Research,* 59, 172–186.

Wilson, K., & French, C.C. (2006). The relationship between susceptibility to false memories, dissociativity, and paranormal belief and experience. *Personality and Individual Differences*, 41, 1493-1502.

Wiseman, R. & Greening, E. (2005). 'It's still bending': Verbal suggestion and alleged psychokinetic metal bending. *British Journal of Psychology*, 96, 115-127.

Wiseman, R., Greening, E., & Smith, M. (2003). Belief in the paranormal and suggestion in the séance room. *British Journal of Psychology*, 94, 285-297.

Wolfradt, U. (1997). Dissociative experiences, trait anxiety, and paranormal beliefs. *Personality and Individual Differences,* 23, 15–19.

European Journal of Parapsychology
Volume 23.2, pages 189-205

ISSN: 0168-7263

Book Review

Matt Colborn
Independent Researcher, UK

An Essay Review of "Irreducible mind: Towards a psychology for the 21st Century" Edited by Edward F. Kelly and Emily Williams Kelly (2006)

Irreducible mind: Towards a psychology for the 21st century is a radical book for an epoch of extreme neurological reductionism. The causes of behaviours, personality and consciousness are increasingly sought in localised brain areas and genes, and pretty much nowhere else; 'there is nothing more, no magic, no additional components to account for every thought, each perception and emotion, all our memories, our personality, fears, loves and curiosities. (O'Shea, 2008, p. 12).' Accompanying this is a widespread attempt to erase views of the mind that are 'occult' or 'immeasurable' (Dennett, 1978; Rorty 1979).

The authors of Irreducible Mind (henceforth IM) believe that there is more than enough evidence to demonstrate that this dominant view is at best incomplete and at worst false. The central problem is the relation between the private subjective 'first person' inner world and the outer, objective 'third person' world of physiological events and processes in the body and brain (p. xvii, E.F. Kelly).

A far more inclusive theory is also needed to accommodate the full range of human experiences, which includes 'rogue phenomena' like psi effects, mind-body interaction, mystic experience and others. F.W.H. Myers sketched the foundations of such a theory over a century ago in his *Human Personality and its Survival of Bodily Death* (1902), which is included on an accompanying CD. In *Human Personality*, Myers developed a theory of the 'Subliminal Self.' This was the idea that:

Correspondence details: Matt Colborn, 14 Hall Road, Haconby, Bourne, Lincolnshire, PE10 0UY, United Kingdom. Email: Matt.Colborn@googlemail.com

> There exists a more comprehensive consciousness… from which the consciousness and faculty of earth-life are mere selections…. [N]o Self of which we can here have cognisance is in reality more than a fragment of a larger Self – revealed in a fashion at once shifting and limited through an organism not so framed as to afford it full manifestation. (Myers, 1902, v2, pp12-15. Quoted in IM p.73)

Myers' Subliminal Mind contained the potential for genius as well as pathology and the emergence of what he termed 'supernormal phenomena' – things that apparently belonged to a more advanced stage of evolution (Myers, 1885). In addition, he thought that the brain should be viewed as a 'filter' that limited consciousness rather than its producer. Myers' theory is therefore strongly dualistic. Most contemporary researchers reject any form of dualism out of hand (Dennett, 1991), but this has resulted in the toleration of some very unsatisfactory models of mind (Searle, 1992).

IM in part presents the case for some kind of interactive dualism. It begins with an overview of contemporary cognitive neuroscience (Chapter 1) and the historical foundations of an alternative approach (chapter 2) before tackling problems of the mind's influence on the body (chapter 3), memory (chapter 4), automatism and secondary centres of consciousness (chapter 5), unusual experiences near death (chapter 6), genius (chapter 7) and mystical experience (chapter 8). Chapter 9 is a re-assessment of Myers' theory and an attempt to provide a sketch of an 'expanded psychology'. For clarity, the individual chapter reviews are subtitled.

Chapter 1: A view from the mainstream: contemporary cognitive neuroscience and consciousness debates. [E.F. Kelly].

This chapter provides a potted history of cognitive neuroscience over the twentieth century, starting with behaviourism and proceeding to the cognitive revolution. This "can be characterized… as a movement toward progressively less unsatisfactory analyses of the mind" (p. 2).

Behaviourism ignored the inner workings of the mind in favour of observable behaviour; complex behaviours were believed to be combinations of simpler stimulus-response reactions. The behaviourist programme gave way in the 1950s and 1960s to a family of positions known collectively as identity theory. Identity theory states that associated mental and physical states are in some sense identical (Feigl, 1958; Putnam, 1967). Putnam (1967) favoured a variety of identity theory called functionalism, which is promoted by contemporary writers like Dennett (1991). Functionalism defines mental states not in terms of what they are made of, but what they do. Just as 'cutting tools' might be made of stone, metal or laser beams, so mental states might be realizable in either brains or computers provided they serve similar functions.

Computers were central to the cognitive revolution because of the idea that a sophisticated computer or Artifical Intelligence (AI) might perform cognitive functions comparable to that of people (Turing, 1950). Some philosophers rejected the idea that a digital computer could ever be conscious. Searle (1992, 1980) argued that since computation was just symbol shuffling, even a sophisticated program could never, for instance, understand language in the way humans do. However, Searle (1984) also advocates a form of 'biological naturalism' (brains = mind), a stance Kelly rejects. I personally find Searle's arguments inconclusive, and they are of course regularly pilloried by writers of a more computationalist bent (e.g. Hofstader, 2007).

The 'second' cognitive revolution involved connectionism and dynamic systems, both of which try to address the inadequacies of classical AI. These newer approaches still rely on simulation on a digital computer, with its ultimate limitations.

There is currently significant dissatisfaction with dominant theories of mind, which belies the triumphal claims of writers like Pinker (2004). Chomsky (1993), a pioneer of the cognitive revolution, complained that it was premature to try and reduce psychology to present-day neurophysiology, and others have made similar noises. Despite the hype, there is a clear need to move beyond the mainstream to develop new mind-models.

Chapter 2: F.W.H Myers and the empirical study of the mind-body problem [E.W.Kelly]

George Santayana's comment (1905) about those who forget history being doomed to repeat it is certainly true of psychology. Many of the current claims that the 'self,' 'free will,' 'qualia,' etc. are illusions are actually just the latest versions of assertions made in the nineteenth century.

In the latter half of the nineteenth century, psychology was transformed from a branch of moral philosophy to a naturalistic science. At the time, mechanistic determinism – the notion of the Universe as machine that obeyed laws -- was the dominant paradigm in science. A result of this was the emergence of epiphenomenalism, the view that mind and free will are illusions or at best secondary, ineffectual by-products of the physical, mechanical world. The idea that the mind might be a fundamental, causal agent in the universe was rejected from the start because this was seen as a primitive, 'supernatural' way of thinking.

However, this rejection had a cost, which was a conflict between experience and knowledge. Individual first-person experience suggests one kind of world of personal agency (free will, volition), but science suggested another kind of world, that of impersonal agency only. This conflict led to a dilemma; either psychology should be expanded to include subjective agency, or else they could narrow psychology to fit science as it was then understood. Historically, the mainstream has chosen the latter course.

Myers remained committed to the idea that the mind might be a fundamental causal agent. This alone was unpalatable to many contemporary scientists, and would be even more so today. However, Myers had many important insights into the nature of consciousness, which are based on many, many case studies. In many ways, he was the Darwin of psychology, in that he was able to organise and classify many different but overlapping phenomena into a coherent whole. The latter half of this important chapter gives an overview of this classification.

Chapter 3: Psychophysiological influence [E.W.Kelly]

The influence of the mind on the body's health, or otherwise, is one of the great unknowns of medicine. This has been denied, of course; the currently prevailing view is that 'the mind-body problem… cannot be viewed as the subject matter of psychosomatic medicine (Lipowski, 1984)! However, there are a wide range of phenomena that beg important questions about the relationship between the body and the mind. 'Psychophysiological influence' refers to a large range of phenomena, from the relatively conventional 'placebo effects' to the hypnotic induction of wounds and marks on the skin (Moody, 1946, 1948), to very odd cases where previous lives seem to have left their mark on the claimant's body (Stevenson, 1997).

The assessment of even relatively conventional phenomena like the placebo effect is not easy. Collins and Pinch (2005) describe placebo research as a 'hall of mirrors,' plagued with expectancy effects and reporting biases. Similar problems emerge all over this field. One running theme is that phenomena tend to get ignored until someone conceives of a (conventional) mechanism to explain them. An example of this is the curing of warts by suggestion, which is a relatively well-attested occurrence. Most of the theories have proposed the alteration of the blood-supply to the warts. Kelly argues that these explanations must be incomplete because some cases involve, for example, the selective reduction of warts.

As an outsider to these fields, I find it very difficult to assess just how well established many of these phenomena really are. For example, Buckman (2007) claimed that there was little persistent evidence that mental attitude affected the course of cancer. Kelly cites several small studies by Meares (1977, 1979,1980) that seem to show that some cancer patients did benefit from an intensive course of meditation – including five that had a complete remission. However Coyne, Stefanek & Palmer (2007) recently published a study on head cancer in which they claimed that psychotherapy had no effect. Emily Kelly (personal communication) responded that there are many factors involved, and ones that are not likely to be reproduced easily in studies like Coyne, Stefanek & Palmer.

Despite these issues, if even only a small fraction of the cases cited by Kelly are legitimate, then they show that our understanding of the interrelationship between the mind and the body is not complete. So although this field has problems, it is a very rich one that deserves far more study than it gets.

Chapter 4: Memory [A.Gauld]

Alan Gauld's discussion of memory and the brain covers firstly problems with what he refers to as the 'memory trace doctrine' and secondly, issues relating to memory and the question of Survival of bodily death.

> There are several serious problems with the idea that our memories are stored as physical 'traces' in the brain. William James' noted that the simple revival of an image into the mind is not enough to make it a memory. Such an image is 'a duplicate, a second event, having absolutely no connection with the first event except that it happens to resemble it….' (James, 1890, pp 649—50, cited in *IM* p.243).

Many explanations of the revival of a past memory presuppose its function and cannot really explain memory itself. Other issues include the richness and form of associations with a particular memory, the literal nature or otherwise of 'mental images' and the feeling of past-directness of personal memories. The latter dilemma is connected to the problem of intentional recall, and general issues of Intentionally. These issues are difficult, but Gauld's prose is precise and deserves careful attention.

Contemporary research into memory tends to be technically sophisticated but philosophically naïve. The Computational Theory of Mind (CTM) assumes that memory works in a way somewhat comparable to the information processing of computers. Many psychologists still discuss memory in terms of 'input' 'encoded' by successive stages of sensory pathways and later 'retrieved' and further processed. (Gauld regards neural network and dynamic systems models of memory as on the same footing as conventional digital-computer models in that their discrete internal memory states are still representations of a kind). Many of the conceptual issues remain.

Contemporary neuroscience is currently dominated by NeuroImaging techniques like PET and fMRI scans. Rather than providing clear answers to issues of memory, fMRI studies have produced rich cohorts of new data, much complexity and many new questions. These studies are also fraught with conceptual and methodological problems including poor to nonexistent replication (Harpaz, 1997). Gauld also notes that the early studies in particular led to some 'breathtaking oversimplifications;' like claims of the 'God spot' or 'seat of intelligence (p.267).'

Gauld concludes that many of these difficulties suggests that 'a fuller understanding of human memory may ultimately require some radical change of perspective.' (p. 281), especially if we take account of the evidence for the survival of memories beyond death. Any theory will be need to be"top down" rather than "bottom up," will accommodate the fact that some declarative memories cannot be tied to localities in the brain, and will regard nerve tracts not as conveyers of information but as 'means by which spatiotemporal patterns of activity in different regions may be fine-tuned to create overarching patterns (p. 228).' Each of these points seems pretty reasonable to me, and would begin address the naiveté of much current memory research.

Chapter 5: Automatism and secondary centres of consciousness [Adam Crabtree]

'Automatism' refers to the occurrence in some people of 'a set of memories, thoughts and feelings which are extra-marginal and outside of the primary consciousness altogether (James 1902/1958. P. 188, cited in IM, p.302). Dissociative Identity Disorder (DID, formerly known as Multiple Personality Disorder), where a number of distinct personalities appear to coexist within one person, is probably the most familiar of these conditions, but there are a whole spectrum of 'dissociative' phenomena that can also be classed as automatisms.

These were crucial for the development of Myers' theory of the 'subliminal mind,' and he thought that the explanation for automatisms needed to be psychological rather than physiological. This was partly because the 'secondary centres of consciousness,' of which the subject was apparently not consciously aware, often appeared to exhibit high levels of volition and intelligence. Examples are the messages sent by 'communicators' through automatic writing and also that some 'secondary personalities' of DID suffers can appear more intelligent and mature than their primary personalities.

The chapter goes on to discuss a number of other historical theories that were held before the subject of automatisms was, like so many other psychological phenomena, pushed into obscurity by the rise of behaviourism. It's somewhat disheartening to read that more recent theorists like Kihlstrom (1993) seem only just to be reaching conclusions that were reached by James and Myers a century ago.

Volition sits at the heart of the contemporary debates on automatisms, and these phenomena have been used to assert that volition or free will is an illusion. Wegner (2002) rejects entirely the idea that human beings are free agents, claiming that all behaviour is involuntary. In his view, conscious will is simply an illusion.

In this interpretation, automatisms are simply behaviours that occur without the illusory feeling that it's produced by conscious will. Kelly observes that this ignores the close association of automatisms with supernormal phenomena, creativity and genius. I'm not sure how far this omission is relevant to the volition issue; one could just as easily argue that works of genius or supernormal perception are as involuntary as the products of pathology. However, it is another example of how researchers tend to cherry pick phenomena that serve their theoretical agendas, and ignore pieces that may prove inconvenient.

The conflicting views on automatisms are interesting because they reveal the ambiguities in much of the data, and how this can be used to serve several contrasting worldviews. It's also sobering to know how little progress has been made in this field the last century because of scientific prejudice.

Chapter 6: Unusual experiences near death and related phenomena [E.W. Kelly, B. Greyson & E.F. Kelly]

Chapter 6 primarily concerns Near Death Experiences, but also covers Out of Body Experiences, Death Bed Visions, Veridical and Collective Apparitions. Much of what's covered will be familiar to those who've followed NDE debates over the years. The basic phenomenology of NDEs is described, and conventional neurophysiological explanations considered. The authors conclude that whilst each of these explanations has some merit, none can explain the NDE as a whole.

The authors then consider those aspects of the NDE that seem to defy conventional models. These include enhanced mentation, veridical Out-of-Body perceptions and visions of deceased acquaintances. Enhanced mentation refers to the report that 'full-fledged mentation, either normal or even enhanced mental activity (p. 386)' occurs when according to conventional theory, mental activity should be receding. The writers acknowledge the difficulty of assessing such subjective reports, but are impressed by their frequency. Veridical OBE perceptions refer to the experiencer seeing or hearing things that they should have been inaccessible during the period of unconsciousness.

Veridical impressions are also hard to verify, first because well-attested cases are admittedly rare, or according to some authors, probably non-existent (Blackmore, 1993). Second, it's hard to eliminate the possibility that unconscious person might have seen or heard something that was later incorporated into the NDE account, that unwittingly lends it veridicality (Blackmore, 1993).

However, there are some aspects of some NDEs that do suggest something strange is going on. Firstly, a number of researchers have found that the memory of events happening just before or after consciousness loss are often confused or absent, which contrasts with the 'enhanced mentation' reports mentioned above. Second, there are rare cases where events were recalled that were occurring when the experiencer's EEG readings were flat. This is potentially very problematic for conventional neuroscience theories of a 'global workspace' because such theories say that normal mental functioning requires synchronous EEG oscillations in the brain. Thinking seems to be happening when these theories say it is impossible.

NDE research suffers because it's often popularly framed as a dispute between those who 'believe' in the afterlife interpretation and those who think that conventional scientific explanations are sufficient (Kellehear, 1995). This is unfortunate, because such polarized thinking might very well mask genuine anomalies that point to unknown mind/ brain functions. However, more and better well-attested cases are still needed, even if they are rare and difficult to find.

Chapter 7: Genius [E.F. Kelly & M. Grosso]

In many senses the products of genius are the most problematic for mainstream psychology, because they are poorly understood but relied upon for scientific breakthroughs. Myers' view of genius was that it 'should be regarded as a power of utilizing a wider range than other men can utilize of faculties…innate in all…. (Human Personality. Vol. 1, p. 71).' He spoke of a 'Subliminal uprush,' that is the manifestation of strength and concentration of some inward unifying control into the conscious mind. This 'uprush' is often achieved through various forms of automatism and even hallucination.

This view is compatible with Hutchinson's (1939) description of the stages of creativity (preparation, incubation, illumination and verification). However, there is a modern school that that sees genius as an incremental application of ordinary consciousness and denies it's quantitatively different at all (Weisberg, 1999; Perkins, 2000). Kelly and Grosso rightly comment that this is an especially egregious example of a willful disregard of evidence that creativity is actually a very mysterious and complex phenomenon. The incommensurability of creative thought with conscious, logical thought is an important example of where computational models fall down in describing the creative process. Myers was right to highlight that many of the creative aspects of genius lie beyond speech, in an almost mystical realm. The authors also highlight the limits of associationism and neural network views in this respect.

Kelly & Grosso are also right to raise the idea that genius is somehow a personality in transformation, with strong links to the mystical. They summarise four areas in which progress needs to be made. Firstly as with other topics, an expansion of current psychology to accommodate data is needed, second is to address the problem of intentionality, third is to focus on the transformative nature of genius and lastly recognition of the links between mysticism and genius.

Chapter 8: Mystical Experience [E.F. Kelly & M. Grosso].

This chapter is in some ways the heart of the book, because it describes experiences that on occasion seem to show that 'pure, unitary, undifferentiated, self-reflexive consciousness' can exist (p.573). This prima facie contradicts functionalist and computationalist theories of the human mind because it suggests that consciousness has a 'central and even supreme reality' over and above function.

The authors begin by providing an overview of the phenomenology of mystical experience. These include ineffability (that these experiences describe expression), that they seem to be in some sense states of non or trans-verbal knowledge or insight, that they are intense and not often sustained and that the experiences often exhibit certitude as to its importance. To quote James (1902), p. 321); 'In mystic states we both become one with the absolute and we become aware of our oneness.'

A key point is whether persons who have had one or more mystical experiences are really talking about the same thing, or whether they are really talking about distinct experiences that have been moulded by their own cultures and only afterwards interpreted as being similar. Constructivists such as Katz (1978) argued that since any experience is filtered by culture, one cannot really talk about a common state of consciousness found across all cultures. However, separate mystical traditions in very different seem to be very similar and constructivism cannot account for the commonality in the reports.

The problem of 'objective significance,' remains. The authors list several circumstantial arguments that mystics do, in fact, make contact with reality in novel ways. There is the persistent connection between mystical experience and genius. There are its connections with automatism, also associated with creativity and genius. There are the apparent connections with supernormal phenomena. But while certainly suggestive, none of these features provide direct evidence that mystical states are anything other than a novel function of the brain.

However, current neurobiological theories of mystical states are not very good. The authors note that researchers commonly (1) fail to come to grips with the full-blown phenomenology of the mystical experience, (2) often have very poor supporting data and (3) are excessively willing to spin out elaborate neurobiological just-so stories. A case in point is the claimed association between mystical states and Temporal Lobe Epilepsy (Saver & Rabin, 1997), for which the authors report 'no credible evidence (p. 534).'

The latter part of the chapter concern the induction of mystical experiences by drugs, including LSD and Ketamine. Again, the associations of these states with other anomalous experiences like NDE have been exaggerated, but there's enough common ground to see these experiences as related but not identical. The authors also comment on opportunities for further research, which are hampered by the current legal restrictions on drugs. An important problem with the assessment of mystical states is that mainstream neuroscience still regards introspection as a dangerous way to gather evidence, especially for non-ordinary states of consciousness. It's all too easy to dismiss these accounts as not really showing what they purport they show. However, the study of such experiences is important if only because it forces us to confront an important dilemma for psychology, which is; how do we make an objective science of something that can only be directly accessed by subjective means?

Chapter 9: Toward a psychology for the 21st Century [E.F. Kelly]

The final chapter begins with a reassessment of Myers' theory and assesses its implications for future psychology. It begins with a brief look at contemporary (1900s) reviews of Human Personality, which included concerns with the definition of 'Subliminal Mind' and with the inclusion of 'rogue phenomena.'

Several issues already covered are also raised, notably Myers' awareness that psychology must be expanded to accommodate territory previously held to belong to religion or metaphysics. Especially important is Myers' plea for a flexible approach to research, and the need to adapt research methods to new situations rather than trying to fit new problems into existing methods.

Kelly concludes that the current status of Myers' Subliminal self is somewhat mixed. He thinks that the conception remains counterintuitive and logically difficult, and that the supporting evidence is 'less than compelling.' (p.594). However, Kelly regards it as 'definitely possible and perhaps even probable, especially in the light of…mystical experiences' (p.595) that some sort of transmission theory is correct. If this is so, then Survival after death, Myers' central concern, becomes at least a possibility.

Can this view of the human psyche be reconciled with contemporary science? One of the strengths of the theory is that it connects, albeit in general terms, both 'normal' and 'supernormal' phenomena. For example, it makes explicit the connection between psi phenomena and dreaming, genius and mysticism. Kelly also claims that the theory has some predictive value; Myers seems to have anticipated NDEs, and his concept of a 'permeable' boundary between the subliminal and normal consciousness predates the work of Thalbourne (1998) and Hartmann (1991) by a century.

The authors of IM believe that the evidence they present is 'sufficient to falsify all forms of biological naturalism, the current physicalist consensus on mind-brain relations. 'The mind is “irreducible”…. There is at least a fundamental bifurcation in nature that cannot be accounted for in [conventional physicalist terms] and we seem driven toward some sort of animist or pluralist alternative.' (p.605).

Despite the cohorts of data presented, I remain uneasy with this conclusion, for both evidential and conceptual reasons. First, even allowing for the length of the book, there are some significant omissions in the evidence. The relative neglect of experimental parapsychological research is strange given that the authors think that it's crucial in their arguments. In fairness, the psi debates are referenced in an annotated bibliography of psychical research, but more discussion of the results, and their interpretation in terms of Myers' Subliminal Mind would have helped.

One consequence of making the data from psychical research central to the arguments for the mind being 'irreducible' is the importation of a number of practical and conceptual problems from that field. Reading IM, one is sometimes left with the impression that these 'rogue phenomena' are accessible to anyone who has eyes to see. This is certainly not the case with psi results, which are often weak, intermittent and context dependent (Hansen, 2001). Experimenter effects are rampant in this field, and 'decline effects' may also be significant (Colborn, 2007). The reports of many of the more outré phenomena like 'maternal impressions' are remote in time and space, and might also be declining (Stevenson, 1997). Any model of an alternate picture of mind must convincingly accommodate these issues, and it's not clear that this has been done.

There are also sociological problems to bridge. One needs also to acknowledge that academe plays an active part in what Weber termed 'the disenchantment of the world' (Hansen, 2001). This concept shows that the erasure of any magical notion of the mind is and probably will remain, central to any future mainstream psychology. Many of the phenomena presented in IM (Stigmata, Reincarnation, Mystical experience) are just too closely associated with religious or mystical views of the world to ever really be part of the rational academic universe.

Another problem is data ambiguity. Anyone carefully reviewing the data of psychical research will find that quite often the data can be made to cut both ways (Bauer, 2001; Collins & Pinch, 1982), as advocates and counter-advocates interpret the same data in different ways according to their a priori expectations. This issue is not just salient to the existence or otherwise of a phenomenon; quite often we cannot be sure whether the data really points to dualism or not. Laboratory psi phenomena might well indicate that some form of dualism is necessary, but then again they might not. Godbey (1978) argued fairly coherently that psi was not incompatible with central state materialism. It is quite possible to build a coherent case for dualism, and a coherent case against, using much of the same data.

Differences in interpretation often arise because of a difference in a priori expectations about the world, and many of these are not directly testable. Metaphysics is the 'elephant in the room' of consciousness studies. As Midgley (2001) points out, many current discussions about consciousness are really about clashing metaphysics.

This denial of metaphysics is also evident in IM. Whilst the authors (I believe correctly) reject the simplistic assumptions of the mainstream, they fail to make sufficiently explicit that often the argument is not over empirical evidence but over ideological and metaphysical difference. This is not to deny that the cumulative evidence presented points to a far richer, more complex and interesting model of mind than is currently advocated by the mainstream. But the central issue of subjective experience (as outlined on pp. 24—29) remains relatively untouched, and perhaps untouchable (see Nagel, 1974).

Finally, the authors examine two alternative models of mind, which they consider congruent with the evidence presented. The first is a non-Cartesian dualist-interactionist model, and the second a neutral-monist model. I remain suspicious of dualist models for reasons that have nothing to do with the 'denial' of rogue phenomena. Dualism was not only abandoned because of prejudice; the Cartesian split was quite artificial and arbitrary, introduced because Descartes was too honest to ignore the strangeness of the subjective viewpoint, but also unable to relate it to the mechanistic worldview he was developing (Midgley, 2001). My own take is that dualism is in many respects an 'observer effect' of the mind looking at itself and to take it too literally is like looking in the mirror and mistaking the reflection for a second person. In my view, it's fatal to lose sight of its artificial nature. For this (metaphysical?) reason, I find the Kelly's suggestion of a neutral-monist model of mind preferable to a return to a 'substance-dualist' solution. As they rightly state, since quantum theory, 'mind' and 'matter' have been redefined out of existence; neither are quite what we thought they were. Maybe 'promissory materialism' will also eventually redefine itself out of existence.

The work of Henry Stapp (2005) seems especially promising here, which seems to answer many of the criticisms of earlier Quantum models and provide testable hypotheses of psychological phenomena like attention. Stapp provides strong arguments to suggest that the behaviour of low level brain constituents are saturated with quantum effects, which puts the burden of proof on those who deny, not affirm the relevance of quantum theory to brain science. These developments do not mean the gap between theory and psychical research has been fully bridged. Even physicists sympathetic to psi phenomena think that current physics would need to be expanded to accommodate them (Carr, 2008). This gap between theory and evidence is an important obstacle to their case (Stent, 1972).

Despite these weaknesses, IM is an important work for a number of reasons. Firstly, it rehabilitates the writings of Myers, which has languished in obscurity for far too long. Second, it places 'psi' and other 'rogue' phenomena in a wider psychological context, suggesting ways that such findings might be integrated with mainstream knowledge. Third, it provides a careful, thorough and systematic critique of the weaknesses of mainstream claims. Fourth, it lays bear the extent of the denial of pretty much every interesting psychological phenomenon by mainstream science, mostly in the name of demystification. For these reasons, it's a key work for parapsychologists, and deserves also to be widely read by mainstream cognitive scientists.

Publication Details: *Kelly, E.F., Kelly, E.W., Crabtree, A., Gauld, A., Grosso, M. & Greyson, B. (2007) Irreducible Mind: Toward a psychology for the 21st century. Rowman and Littlefield. 800 pp. ISBN-13: 978-0-7425-4792-6, ISBN-10: 0-7425-4792-2. Publication price: $79.95 (Cloth).*

References

Bauer, H. (2001). *Science or pseudoscience: Magnetic healing, psychic phenomena and other heterodoxies*. University of Illinois Press: Urbana and Chicago.

Blackmore, S.J. (1993). *Dying to live*. London: Grafton.

Buckman, R. (1996). *What you really need to know about cancer*. London: Pan Macmillan.

Carr. B. (2008). Worlds Apart? Can Psychical Research Bridge the Gulf Between Matter and Mind? *Proceedings of the society for Psychical Research*, 59, 1-96.
Chomsky, N. (1993). *Language and Thought*. Wakefield, RI: Moyer Bell.
Colborn, M.L.C. (2007). The decline effect in spontaneous and experimental psychical research. *Journal of the Society for Psychical Research*, 71, 1-22.
Collins, H.M. & Pinch, T.J. (2005). *Dr. Golem: How to think about medicine*. Chicago: University of Chicago Press.
Collins, H.M. & Pinch, T.J. (1982). *Frames of meaning: the social construction of extraordinary science*. London: Routledge & Kegan Paul.
Coyne, J.C., Stefanek, M. and Palmer, S.C. (2007). Psychotherapy and survival in cancer: The conflict between hope and evidence. *Psychological Bulletin*, 133, 367-394.
Dennett, D.C. (1991). *Consciousness explained*. Boston: Little, Brown.
Dennett, D.C. (1978). Towards a cognitive theory of consciousness. In *Brainstorms: Philosophical essays in mind and psychology*. Montgomery: Bradford Books.
Feigl, H. (1958). The "mental" and the "physical." In H.Feigl, M.Scriven, & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. 2, pp. 370-497). Minneapolis: University of Minnesota Press.
Gauld, A. (1982). *Mediumship and Survival*. London: Heinemann.
Godbey, J.W. (1978). Central state materialism and parapsychology. In J. Ludwin (Ed.) *Philosophy and parapsychology*. Prometheus books: Buffalo.
Hansen, G. P. (2001). *The Trickster and the Paranormal*. Xlibris.
Harpav, Y. (1997). *Replicability of cognitive imaging of the cerebral cortex by PET and fMRI: A survey of recent literature*. http://human-brain.org/replicability.html, accessed 31/07/08.
Hartmann, E. (1991). *Boundaries in the mind: A new psychology of personality*. New York: Basic Books.
Hofstader, D.R. (2007). *I am a strange loop*. London: Basic Books.
Hutchinson E.D. (1939). Varieties of Insight in Humans. *Psychiatry*, 2, 323-332.
James, W. (1958). *The varieties of religious experience*. New York: Mentor. (Original work published 1902).
James, W. (1890). *The Principles of psychology* (2 Vols). New York: Henry Holt.
Katz,S.T. (Ed.) (1978). *Mysticism and philosophical analysis*. New York: Oxford University Press.
Kellehear, A. (1995). Review of Dying to Live: science and the Near-Death Experience by Blackmore, S. (1993), Prometheus: Buffalo. *Journal of Scientific Exploration*, 9, 312-14.
Kihlstrom, J.F. (1993) The psychological unconscious and the self. In G. R. Bock & J. Marsh (Eds.), *Experimental and theoretical studies of consciousness* (pp. 147-156). Chichester, UK: John Wiley & Sons.
Lipowski, Z.J. (1984). What does the word "psychosomatic" really mean? A historical and semantic enquiry. *Psychosomatic Medicine*, 46, 153-171.
Meares, A. (1980). What can a cancer patient expect from intensive meditation? *Australian Family Physician*, 9, 322-325.

Meares, A. (1979). Meditation: A psychological approach to cancer treatment. *The Practitioner*, 222, 119-122.

Meares, A. (1977). Atavistic regression as a factor in the remission of cancer. *Medical Journal of Australia*, 2, 132-133.

Midgley, M. (2001). *Science and Poetry*. London: Routledge & Kegal Paul.

Moody, R.L. (1946). Bodily changes during abreaction. *Lancet*, 1, 934-935.

Moody, R.L. (1948). Bodily changes during abreaction. *Lancet*, 2, 964.

Myers, F.W.H. (1902). *Human Personality and its survival of bodily death*. (2 vols.). London: Longmans Green.

Myers, F.W.H. (1885). Automatic writing – II. *Proceedings of the Society for Psychical Research*, 3, 1-63.

Nagel, T. (1974). What Is it Like to Be a Bat? *Philosophical Review*, 83, 435-50.

O'Shea, M. (2008). *Science made simple: the Brain* (booklet distributed free with the Independent newspaper in association with Oxford University Press).

Perkins, D.N. (2000). *The Eureka Effect: The Art and logic of Breakthrough Thinking*. New York: W.W. Norton.

Pinker, S. (September 7th, 2004). Editorial, *Newsweek*.

Putnam, H. (1967). Psychological predicates. In W.H. Capitan & d.D. merrill (Eds.), *Art, Mind and Religion* (pp. 37-48). Pittsburgh, PA: University of Pittsburgh Press.

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton, N.J.:Princeton University Press.

Santayans, G. (1905). *Life of Reason, Reason in Common Sense*. Scribner's.

Saver, J.L. & Rabin, J. (1997). The neural substrates of religious experience. *Journal of Neuropsychiatry and Clinical Neuroscience*, 9, 498-510.

Searle, J.R. (1992). *The Rediscovery of Mind*. Cambridge, MA: Harvard University Press.

Searle, J. R. (1984). *Minds, brains and science*. Cambridge, MA: Harvard University Press.

Searle, J.R. (1980). Minds, brains and programs. *Behavioural and Brain sciences*, 3, 417-424.

Stapp, H. (2005). Quantum interactive dualism: An alternative to materialism. *Journal of Consciousness Studies*, 12, 43-58.

Stent, G. S. (1972). Prematurity and Uniqueness in Scientific Discovery. *Scientific American*, December, 84-93.

Stevenson, I. (1997). *Reincarnation and Biology: A contribution to the etiology of birthmarks and birth defects*. (2 vols.). New York: Praeger.

Thalbourne, M.A. (1998) Transliminality: Further correlates and a short measure. *Journal of the American Society for Psychical Research*, 92, 402-419.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.

Varela, F.J. (1997) *Sleeping, dreaming and dying*. Somerville, MA: Wisdom.

Weisberg, R.W. (1999). Creativity and Knowledge: A challenge to theories. In R.J. Sternberg (Ed.), *Handbook of Creativity* (pp. 251-272). New York: Cambridge University Press,

Wegner (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.